



## SEQUENCE LISTING

### (1) GENERAL INFORMATION:

- (i) APPLICANT: Takashi OKADO et al.
- (ii) TITLE OF INVENTION: GENE REGULATING AUREOBASIDIN SENSITIVITY  
(AS AMENDED)
- (iii) NUMBER OF SEQUENCES: 50
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: WENDEROTH, LIND & PONACK, L.L.P.
  - (B) STREET: 2033 K Street, N.W., #800
  - (C) CITY: Washington
  - (D) STATE: D.C.
  - (E) COUNTRY: U.S.A.
  - (F) ZIP: 20006
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
  - (B) COMPUTER: IBM Compatible
  - (C) OPERATING SYSTEM: MS-DOS
  - (D) SOFTWARE: wordperfect 5.1
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: 10/777,195
  - (B) FILING DATE: February 13, 2004
- (vii) PRIOR APPLICATION DATA
  - (A) APPLICATION NUMBER: 09/368,431
  - (B) FILING DATE: August 5, 1999
- (viii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/716,873
  - (B) FILING DATE: September 20, 1996
- (ix) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Warren M. Cheek, Jr.
- (B) REGISTRATION NUMBER: 33,367
- (C) REFERENCE/DOCKET NUMBER: 2003-1772/WMC/00001
- (x) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 202-721-8200
  - (B) TELEFAX: 202-721-8250
  - (C) TELEX:
- (2) INFORMATION FOR SEQ ID NO:1:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 4140 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: double
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: Genomic DNA

```

agatctgtgg cttccggttg gctacttgta accaactgat ggtcagatgg atctgccgtc   60
tgttttgatt tgaattttcc ctgctcattc tgattctgtg agaggctgca ttcattatca  120
catctcatac ccggcgccctg cgacttcggt cacctctgcg gtctggcggt taccgggggtc  180
cgtctgagac tcgtcagtca gccattcgag tatgcgaact ctgactttgc tcacctaaga  240
gtttgcacga gatgccgaaa tcctcctcga gtagagtttg caaggcttga accttgggtcc  300
ttgaagcccc aaagtggctc agtagtggga tcgatagtct gggtgttgaa gattttctcc  360
tccaccttac ctatggccgc tggccttctc cacctttcag gctttcaggc accctcggct  420
cggattctgt atcgtccggt accgaagcta gtcctagcta gtcaaagcta gtccaagcta  480
gtctcgtcaa ggtttggcgc agcgcggttc cgtgtaaagt acaaatttga aatacgaata  540
cgcagtactc gcagccggca cttccgctca gcccaggctc agaggctaag ggtgttggcg  600
cttcctcatc atcttcttct cgtcgacctt ttctctttc tctccctatc ggtgcttctc  660
tccaacctca ttctcagtcg ttcgcccatc aggtttatac tccggctccg tggccatctg  720
cctccctcac gacctcctcg ttccaggttt tcctctcgac tgctgcgccc ttgcacttcg  780
ccttgcatca gtgaaacccc ctgcaacgtg acgggtcaaa gacatcctcg tttggccgct  840
ggagaccgga gcgtgcgctt cgtttcgtct tcttcgaacc gatctcaatt tccccgctcg  900
ggttgacgcc gtcagacccc tgctcgttgc ctaacggctt gttattcaag accccttttc  960

```

tgccgcttcc gcgaccgatt tattcgtcgc cttccaactc ttgtacaatc ggggggaaag 1020  
 aaagcagacg gagttcgatc tggaggaatt atagctgagt cttgcccgca agactcgccg 1080  
 caaccatgaa tcaaacactt cccacgtgga aggaccgcac gcagaaccag tttggaaagc 1140  
 ttcagatcca ggttccatgg cgggccatcc aactgctcgt cccgcatcgc atgcggcgga 1200  
 agttaaggtc caaattgcgc agtagagcgt ctcctacctc gtcaatagcc tctttacaga 1260  
 cgtcgttatc gcctgcagac aactacgat cgctccaaag ccaccgatgg acggtttacg 1320  
 acttccaata tctgcttctg ttgatcgtgg gcatcttctc tttgaccgtt atcgagtcgc 1380  
 ccgggccttt gggcaaaacg gccattttct ccatgctcct attctctctc ctgatcccta 1440  
 tgacccgcca gttcttctc ccgtttctgc cgattgccgg atggcttctg tttttctacg 1500  
 cctgccagtg agttaaaaaac aaccgcgtac cagaccccggt gcagcagtta ctcacatatg 1560  
 caggttcatc ccaagcgatt ggcgccctgc gatttgggtt cgtgtcttgc ctgcactgga 1620  
 gaatattctc tacggcgcaa acatcagcaa catcctatcc gctcaccaga acgttggtgct 1680  
 tgacgtgctg gcgtggctac cctacggtat ctgccactat ggcgctccgt ttgtgtgctc 1740  
 gttgatcatg ttcattctcg gtccgcccgg cactgttccc cttttcgcgc gcactttcgg 1800  
 ctatatcagt atgactgcgg ttactattca gctgtttttc ctttgctctc caccttggtgta 1860  
 tgagaatcgc tatggtctag ctccggcaga ctactccatc caaggtgatc ccgcagggct 1920  
 tgcccgcatc gacaagcttt tcggcatcga cctttacacg tctgttttcc atcagtcgcc 1980  
 tgttgtgttc ggcgcttttc cgtcgctgca tgctgccgac tcaaccctgg ccgcactttt 2040  
 catgagtcac gttttccccc gcatgaagcc cgtcttcgtg acctatactc tatggatgtg 2100  
 gtgggcaaca atgtacctc cacatcacta tgcggtcgat ttggttgagg gtggtctcct 2160  
 ggccgccatt gctttctact tcgccaagac ccgattcctt ccccgtgtcc agctcgacaa 2220  
 gaccttccgt tgggactacg actatgtgga attcggcgag tctgccctgg agtatgggta 2280  
 tgggtgcagct ggctatgatg gagacttcaa tctcgacagc gatgaatgga ctggttggtc 2340  
 ttcattcctc gtctcctcag gtccttgag tcccgttgac gatcattact catgggaaac 2400  
 cgaggcactg acctccccac atactgatat tgagtccggc aggcatactt tcagcccttg 2460  
 agtagccaca aaccaaactc gatactgca tatagcgatc tcgctcctcc tccactgcat 2520  
 ctatttacga gacggcggtta gaacatttca cgacattctg gctttattgc atcgagcaca 2580  
 tttcgacaca tatacttta atacccttc ttcggtgtcc cagatcatcg gttcgacctt 2640  
 aatgtacctc ggtccgaatc cgcctgggat actgtttctc tttccgccgc acttcactgt 2700  
 acattgcttg acattgcgaa accgggttgg gctcgaacgt gggatggggtt atcgctcatc 2760  
 gctacacgcc gttgctccat cataatgtta atggacaaa tggggctacg catcctgggtg 2820  
 tttagtcctg gaagaccatc cgataacccc cgtcggtaac actcgcttgt ctcgtgtcca 2880

cccagacact acttcaattc tcacttctat cgtccgctat taccttgacc tggtcgaacc 2940  
 catccttatt attcgtttcg actatgctat atatttattt ttaccattcg tgtcgatcgc 3000  
 tcatactctt ggcgcttggg actggaagca tttatattgg aaaaaatcac ggaatggggc 3060  
 gccttttctt cttgcacttc actcgctgtg catagacggg tttacatttc tgctttgcaa 3120  
 tgcatacaga actctgcatt agcatataga aagaggggaa ggatggacct tcttcttgat 3180  
 tgctcgcag gtttatccat tcgctcaaag tggattacgt ccacatatta cccgggggct 3240  
 atacacatgg ctactgtgtt gctttctgac attcgccgga cgtgcaaggt tgggaggaga 3300  
 gtctgacgct gacggggctt gttgaaggat gttcacgctg cccgatttga cccggcttcg 3360  
 actaacctca gattctcgac ttgttggacg gtgacttgac ttgcttgcta tggctgacg 3420  
 ctctcacacc tacctatcac atcctctca cctcacaaat tccgctcatg gacactatcc 3480  
 tcttcttttc gtttcccttg gatagtgtgt gtgtgtgtgt ggttggggca aattatccat 3540  
 agcagcagta ttattagtta taatccggta gtgttatgat ttatgaaggc aacttgata 3600  
 ctattgccac tttgtccata tctcttgctt gtaatagaac tgacatcgcg acgcttcggc 3660  
 cacgatgcat ataaaaactc tagtcaacac gatattaaca agcgaaacca ttacgctgta 3720  
 aactattcag gatcgccgcg ggcccatctg ggacttgact gtactaaata tgtcctaaag 3780  
 caagcagact aaatatttaa cgtgggatat tattcatata cgcatatgta tacatagtca 3840  
 taacaagcca aggggtgggt aggggtgggt aattattatt ttttttcgtc gatacaagta 3900  
 tccatcctta aatgtccgtg gtctactctt cataaatctt aaccgcgtcc gcatactcct 3960  
 ttatcctcga gacaaaagtg tcttcaattt catcgccacg gccaccagca acgcgaggga 4020  
 taaggtcggt gagagaggcg cggccgagaa tgtcgacatg gtcgccttca tattgttagc 4080  
 atgcgacgtc agactggaac caggaagga aaggagagag gtacctgtat ttggaccacc 4140

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 439 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asn Gln Thr Leu Pro Thr Trp Lys Asp Arg Thr Gln Asn Gln

1	5	10	15
Phe Gly Lys Leu	Gln Ile Gln Val	Pro Trp Arg Ser	Ile Gln Leu
20	25	30	
Leu Val Pro His	Arg Met Arg Arg	Lys Leu Arg Ser	Lys Leu Arg
35	40	45	
Ser Arg Ala Ser	Pro Thr Ser Ser	Ile Ala Ser Leu	Gln Thr Ser
50	55	60	
Leu Ser Pro Ala	Asp Thr Leu Arg	Ser Leu Gln Ser	His Arg Trp
65	70	75	
Thr Val Tyr Asp	Phe Gln Tyr Leu	Leu Leu Ile Val	Gly Ile
80	85	90	
Phe Ser Leu Thr	Val Ile Glu Ser	Pro Gly Pro Leu	Gly Lys Thr
95	100	105	
Ala Ile Phe Ser	Met Leu Leu Phe	Ser Leu Leu Ile	Pro Met Thr
110	115	120	
Arg Gln Phe Phe	Leu Pro Phe Leu	Pro Ile Ala Gly	Trp Leu Leu
125	130	135	
Phe Phe Tyr Ala	Cys Gln Phe Ile	Pro Ser Asp Trp	Arg Pro Ala
140	145	150	
Ile Trp Val Arg	Val Leu Pro Ala	Leu Glu Asn Ile	Leu Tyr Gly
155	160	165	
Ala Asn Ile Ser	Asn Ile Leu Ser	Ala His Gln Asn	Val Val Leu
170	175	180	
Asp Val Leu Ala	Trp Leu Pro Tyr	Gly Ile Cys His	Tyr Gly Ala
185	190	195	
Pro Phe Val Cys	Ser Leu Ile Met	Phe Ile Phe Gly	Pro Pro Gly
200	205	210	
Thr Val Pro Leu	Phe Ala Arg Thr	Phe Gly Tyr Ile	Ser Met Thr
215	220	225	
Ala Val Thr Ile	Gln Leu Phe Phe	Pro Cys Ser Pro	Pro Trp Tyr
230	235	240	
Glu Asn Arg Tyr	Gly Leu Ala Pro	Ala Asp Tyr Ser	Ile Gln Gly

245	250	255
Asp Pro Ala Gly Leu Ala Arg Ile Asp	Lys Leu Phe Gly Ile Asp	
260	265	270
Leu Tyr Thr Ser Val Phe His Gln Ser	Pro Val Val Phe Gly Ala	
275	280	285
Phe Pro Ser Leu His Ala Ala Asp Ser	Thr Leu Ala Ala Leu Phe	
290	295	300
Met Ser His Val Phe Pro Arg Met Lys	Pro Val Phe Val Thr Tyr	
305	310	315
Thr Leu Trp Met Trp Trp Ala Thr Met	Tyr Leu Ser His His Tyr	
320	325	330
Ala Val Asp Leu Val Ala Gly Gly Leu	Leu Ala Ala Ile Ala Phe	
335	340	345
Tyr Phe Ala Lys Thr Arg Phe Leu Pro	Arg Val Gln Leu Asp Lys	
350	355	360
Thr Phe Arg Trp Asp Tyr Asp Tyr Val	Glu Phe Gly Glu Ser Ala	
365	370	375
Leu Glu Tyr Gly Tyr Gly Ala Ala Gly	Tyr Asp Gly Asp Phe Asn	
380	385	390
Leu Asp Ser Asp Glu Trp Thr Val Gly	Ser Ser Ser Ser Val Ser	
395	400	405
Ser Gly Ser Leu Ser Pro Val Asp Asp	His Tyr Ser Trp Glu Thr	
410	415	420
Glu Ala Leu Thr Ser Pro His Thr Asp	Ile Glu Ser Gly Arg His	
425	430	435
Thr Phe Ser Pro		

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2856 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: cDNA  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

ggtttataact ccggctccgt ggccatctgc ctccctcacg acctcctcgt tccaggtttt 60
cctctcgact gctgcgccct tgcacttcgc cttgcatcag tgaaaccccc tgcaacgtga 120
cggctcaaag acatcctcgt ttggccgctg gagaccggag cgtgcgcttc gtttcgtctt 180
cttcgaaccg atctcaatct ccccgctcgg gttgacgccg tcagcaccct gctcgttgcc 240
taacggcttg ttattcaaga ccccttttct gccgcttcgc cgaccgattt attcgtcgcc 300
ttccaactct tgtacaatcg gggggaaaga aagcagacgg agttcgatct ggaggaatta 360
tagctgagtc ttgcccgcaa gactcgccgc aaccatgaat caaacacttc ccacgtggaa 420
ggaccgcacg cagaaccagt ttggaaagct tcagatccag gttccatggc ggtccatcca 480
actgctcgtc ccgcatcgca tgcggcgga gtttaaggctc aaattgcgca gtagagcgtc 540
tcctacctcg tcaatagcct ctttacagac gtcgttatcg cctgcagaca cactacgatc 600
gctccaaagc caccgatgga cggtttacga cttccaatat ctgcttctgt tgatcgtggg 660
catcttctct ttgaccgtta tcgagtcgcc cgggcctttg ggcaaacgg ccattttctc 720
catgctccta ttctctctcc tgatccctat gaccgcagc ttcttctctc cgtttctgcc 780
gattgccgga tggcttctgt ttttctacgc ctgccagttc atcccaagcg attggcgccc 840
tgcgatttgg gttcgtgtct tgcctgcact ggagaatatt ctctacggcg caaacatcag 900
caacatccta tccgctcacc agaacgttgt gcttgacgtg ctggcggtggc taccctacgg 960
tatctgccac tatggcgctc cgtttgtgtg ctcgttgatc atgttcatct tcggtccgcc 1020
cggcactgtt ccccttttcg cgcgcacttt cggctatatc agtatgactg cggttactat 1080
tcagctgttt ttcccttgct ctccaccttg gtatgagaat cgctatggtc tagctccggc 1140
agactactcc atccaagggtg atcccgcagg gcttgcccgc attgacaagc ttttcggcat 1200
cgacctttac acgtctgttt tccatcagtc gcctgttggtg ttcggcgctt ttccgtcgct 1260
gcatgctgcc gactcaaccc tggccgcact tttcatgagt catgttttcc cccgcatgaa 1320
gcccgtcttc gtgacctata ctctatggat gtggtgggca acaatgtacc tctcacatca 1380
ctatgcggtc gatttggttg cgggtggtct cctggccgcc attgctttct acttcgccaa 1440
gacccgattc cttccccgtg tccagctcga caagaccttc cgttgggact acgactatgt 1500
ggaattcggc gagtctgccc tggagtatgg gtatggtgca gctggctatg atggagactt 1560
caatctcgac agcgatgaat ggactgttgg ttcttcatcc tccgtctcct caggctcctt 1620

```

gagtcccgtt gacgatcatt actcatggga aaccgaggca ctgacctccc cacatactga 1680  
 tattgagtcc ggcaggcata ctttcagccc ttgagtagcc acaaaccaaa ctcgatacct 1740  
 gcatatagcg atctcgctcc tcctccactg catctattta cgagacggcg ttagaacatt 1800  
 tcacgacatt ctggccttat tgcacgagc acatttcgac acatatatct ttaataccct 1860  
 ttcttcggtg tcccagatca tcggttcgac cttaatgtac ctcggtccga atccgcctgg 1920  
 gatactgttt ctctttccgc cgcacttcac tgtacattgc ttgacattgc gaaaccgggt 1980  
 tgggctcgaa cgtgggatgg gttatcgctc atcgctacac gccgttgctc catcataatg 2040  
 ttaatggaca caatggggct acgcatcctg gtgttttagtc ctggaagacc atccgataac 2100  
 ccccgctcgg aacactcgct tgtctcgtgt ccaccagac actacttcaa ttctcacttc 2160  
 tatcgctcgc tattaccttg acctggtcga acccatcctt attattcggt tcgactatgc 2220  
 tatatatatta tttttaccat tcgtgtcgat cgctcatact cttggcgctt gggactggaa 2280  
 gcatttatat tggaaaaaat cacggaatgg ggcgcctttt cttcttgac ttcactcgct 2340  
 gtgcatagac gggttttacat ttctgctttg caatgcatca cgaactctgc attagcatat 2400  
 agaaagaggg gaaggatgga ctttcttctt gattgctcgc atgggttatc cattcgctca 2460  
 aagtggatta cgtccacata ttaccggggg gctatacaca tggctactgt gttgctttct 2520  
 gacattcgcc ggacgtgcaa gggtgggagg agagtctgac gctgacgggg cttgttgaag 2580  
 gatgttcacg cgtcccgatt tgaccgggt tcgactaacc tcagattctc gacttgttgg 2640  
 acggtgactt gacttgcttg ctatggcttg acgctctcac acctacctat cacatcctcc 2700  
 tcacctcaca aattccgctc atggacacta tcctcttctt ttcgtttccc ttggatagtg 2760  
 tgtgtgtgtg tgtggttggg gcaaattatc catagcagca gtattattag ttataatccg 2820  
 gtagtgttat gatttatgaa ggcaacttgt atacta 2856

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 439 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asn Gln Thr Leu Pro Thr Trp Lys Asp Arg Thr Gln Asn Gln



1	5	10	15
Phe Gly Lys Leu Gln Ile Gln Val Pro Trp Arg Ser Ile Gln Leu			
	20	25	30
Leu Val Pro His Arg Met Arg Arg Lys Leu Arg Ser Lys Leu Arg			
	35	40	45
Ser Arg Ala Ser Pro Thr Ser Ser Ile Ala Ser Leu Gln Thr Ser			
	50	55	60
Leu Ser Pro Ala Asp Thr Leu Arg Ser Leu Gln Ser His Arg Trp			
	65	70	75
Thr Val Tyr Asp Phe Gln Tyr Leu Leu Leu Leu Ile Val Gly Ile			
	80	85	90
Phe Ser Leu Thr Val Ile Glu Ser Pro Gly Pro Leu Gly Lys Thr			
	95	100	105
Ala Ile Phe Ser Met Leu Leu Phe Ser Leu Leu Ile Pro Met Thr			
	110	115	120
Arg Gln Phe Phe Leu Pro Phe Leu Pro Ile Ala Gly Trp Leu Leu			
	125	130	135
Phe Phe Tyr Ala Cys Gln Phe Ile Pro Ser Asp Trp Arg Pro Ala			
	140	145	150
Ile Trp Val Arg Val Leu Pro Ala Leu Glu Asn Ile Leu Tyr Gly			
	155	160	165
Ala Asn Ile Ser Asn Ile Leu Ser Ala His Gln Asn Val Val Leu			
	170	175	180
Asp Val Leu Ala Trp Leu Pro Tyr Gly Ile Cys His Tyr Gly Ala			
	185	190	195
Pro Phe Val Cys Ser Leu Ile Met Phe Ile Phe Gly Pro Pro Gly			
	200	205	210
Thr Val Pro Leu Phe Ala Arg Thr Phe Gly Tyr Ile Ser Met Thr			
	215	220	225
Ala Val Thr Ile Gln Leu Phe Phe Pro Cys Ser Pro Pro Trp Tyr			
	230	235	240
Glu Asn Arg Tyr Gly Leu Ala Pro Ala Asp Tyr Ser Ile Gln Gly			

245	250	255
Asp Pro Ala Gly Leu Ala Arg Ile Asp Lys Leu Phe Gly Ile Asp		
260	265	270
Leu Tyr Thr Ser Gly Phe His Gln Ser Pro Val Val Phe Gly Ala		
275	280	285
Phe Pro Ser Leu His Ala Ala Asp Ser Thr Leu Ala Ala Leu Phe		
290	295	300
Met Ser His Val Phe Pro Arg Met Lys Pro Val Phe Val Thr Tyr		
305	310	315
Thr Leu Trp Met Trp Trp Ala Thr Met Tyr Leu Ser His His Tyr		
320	325	330
Ala Val Asp Leu Val Ala Gly Gly Leu Leu Ala Ala Ile Ala Phe		
335	340	345
Tyr Phe Ala Lys Thr Arg Phe Leu Pro Arg Val Gln Leu Asp Lys		
350	355	360
Thr Phe Arg Trp Asp Tyr Asp Tyr Val Glu Phe Gly Glu Ser Ala		
365	370	375
Leu Glu Tyr Gly Tyr Gly Ala Ala Gly Tyr Asp Gly Asp Phe Asn		
380	385	390
Leu Asp Ser Asp Glu Trp Thr Val Gly Ser Ser Ser Ser Val Ser		
395	400	405
Ser Gly Ser Leu Ser Pro Val Asp Asp His Tyr Ser Trp Glu Thr		
410	415	420
Glu Ala Leu Thr Ser Pro His Thr Asp Ile Glu Ser Gly Arg His		
425	430	435
Thr Phe Ser Pro		

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 436 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Asn	Thr	Thr	Leu	Pro	Ser	Trp	Lys	Asp	Arg	Thr	Gln	Asn	Gln
1				5				10					15	
Phe	Gly	Lys	Leu	Gln	Ile	Gln	Val	Pro	Trp	Arg	Thr	Ile	Gln	Leu
				20				25					30	
Leu	Val	Pro	His	Arg	Met	Arg	Arg	Lys	Ile	Arg	Ser	Lys	Leu	Arg
				35				40					45	
Ser	Arg	Ile	Ser	Pro	Thr	Ser	Ser	Ile	Ser	Ser	Leu	Gln	Thr	Ser
				50				55					60	
Phe	Ser	Pro	Val	Asp	Thr	Leu	Arg	Ser	Leu	Gln	Ser	His	Arg	Trp
				65				70					75	
Thr	Leu	Tyr	Asp	Phe	Gln	Tyr	Leu	Leu	Leu	Ile	Val	Gly	Ile	
				80				85					90	
Phe	Ser	Leu	Ser	Val	Met	Glu	Ser	Pro	Gly	Pro	Leu	Ala	Lys	Thr
				95				100					105	
Ala	Ala	Phe	Thr	Leu	Leu	Leu	Val	Ser	Leu	Leu	Leu	Pro	Ile	Thr
				110				115					120	
Arg	Gln	Phe	Phe	Leu	Pro	Phe	Leu	Pro	Ile	Ala	Gly	Trp	Leu	Ile
				125				130					135	
Phe	Phe	Tyr	Ala	Cys	Gln	Phe	Ile	Pro	Ser	Asp	Trp	Arg	Pro	Ala
				140				145					150	
Ile	Trp	Val	Arg	Val	Leu	Pro	Ala	Leu	Glu	Asn	Ile	Leu	Tyr	Gly
				155				160					165	
Ala	Asn	Ile	Ser	Asn	Ile	Leu	Ser	Ala	His	Gln	Asn	Val	Val	Leu
				170				175					180	
Asp	Val	Leu	Ala	Trp	Leu	Pro	Tyr	Gly	Ile	Cys	His	Tyr	Gly	Ala
				185				190					195	
Pro	Phe	Val	Cys	Ser	Ala	Ile	Met	Phe	Ile	Phe	Gly	Pro	Pro	Gly
				200				205					210	

Thr Val Pro Leu Phe Ala Arg Thr Phe Gly Tyr Ile Ser Met Ala		
	215	220 225
Ala Val Thr Ile Gln Leu Phe Phe Pro Cys Ser Pro Pro Trp Tyr		
	230	235 240
Glu Asn Leu Tyr Gly Leu Ala Pro Ala Asp Tyr Ser Met Pro Gly		
	245	250 255
Asn Pro Ala Gly Leu Ala Arg Ile Asp Glu Leu Phe Gly Ile Asp		
	260	265 270
Leu Tyr Thr Ser Gly Phe Arg Gln Ser Pro Val Val Phe Gly Ala		
	275	280 285
Phe Pro Ser Leu His Ala Ala Asp Ser Thr Leu Ala Ala Leu Phe		
	290	295 300
Met Ser Gln Val Phe Pro Arg Leu Lys Pro Leu Phe Val Ile Tyr		
	305	310 315
Thr Leu Trp Met Trp Trp Ala Thr Met Tyr Leu Ser His His Tyr		
	320	325 330
Ala Val Asp Leu Val Gly Gly Gly Leu Leu Ala Thr Val Ala Phe		
	335	340 345
Tyr Phe Ala Lys Thr Arg Phe Met Pro Arg Val Gln Asn Asp Lys		
	350	355 360
Met Phe Arg Trp Asp Tyr Asp Tyr Val Glu Tyr Gly Asp Ser Ala		
	365	370 375
Leu Asp Tyr Gly Tyr Gly Pro Ala Ser Phe Glu Gly Glu Phe Asn		
	380	385 390
Leu Asp Ser Asp Glu Trp Thr Val Gly Ser Ser Ser Ser Ile Ser		
	395	400 405
Ser Gly Ser Leu Ser Pro Val Asp Asp His Tyr Ser Trp Glu Gly		
	410	415 420
Glu Thr Leu Ala Ser Pro Ala Thr Asp Ile Glu Ser Gly Arg His		
	425	430 435
Phe		

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 401 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Ala	Asn	Pro	Phe	Ser	Arg	Trp	Phe	Leu	Ser	Glu	Arg	Pro	Pro
1				5					10					15
Asn	Cys	His	Val	Ala	Asp	Leu	Glu	Thr	Ser	Leu	Asp	Pro	His	Gln
				20					25					30
Thr	Leu	Leu	Lys	Val	Gln	Lys	Tyr	Lys	Pro	Ala	Leu	Ser	Asp	Trp
				35					40					45
Val	His	Tyr	Ile	Phe	Leu	Gly	Ser	Ile	Met	Leu	Phe	Val	Phe	Ile
				50					55					60
Thr	Asn	Pro	Ala	Pro	Trp	Ile	Phe	Lys	Ile	Leu	Phe	Tyr	Cys	Phe
				65					70					75
Leu	Gly	Thr	Leu	Phe	Ile	Ile	Pro	Ala	Thr	Ser	Gln	Phe	Phe	Phe
				80					85					90
Asn	Ala	Leu	Pro	Ile	Leu	Thr	Trp	Val	Ala	Leu	Tyr	Phe	Thr	Ser
				95					100					105
Ser	Tyr	Phe	Pro	Asp	Asp	Arg	Arg	Pro	Pro	Ile	Thr	Val	Lys	Val
				110					115					120
Leu	Pro	Ala	Val	Glu	Thr	Ile	Leu	Tyr	Gly	Asp	Asn	Leu	Ser	Asp
				125					130					135
Ile	Leu	Ala	Thr	Ser	Thr	Asn	Ser	Phe	Leu	Asp	Ile	Leu	Ala	Trp
				140					145					150
Leu	Pro	Tyr	Gly	Leu	Phe	His	Phe	Gly	Ala	Pro	Phe	Val	Val	Ala
				155					160					165
Ala	Ile	Leu	Phe	Val	Phe	Gly	Pro	Pro	Thr	Val	Leu	Gln	Gly	Tyr

170	175	180
Ala Phe Ala Phe Gly Tyr Met Asn Leu	Phe Gly Val Ile Met Gln	
185	190	195
Asn Val Phe Pro Ala Ala Pro Pro Trp	Tyr Lys Ile Leu Tyr Gly	
200	205	210
Leu Gln Ser Ala Asn Tyr Asp Met His	Gly Ser Pro Gly Gly Leu	
215	220	225
Ala Arg Ile Asp Lys Leu Leu Gly Ile	Asn Met Tyr Thr Thr Ala	
230	235	240
Phe Ser Asn Ser Ser Val Ile Phe Gly	Ala Phe Pro Ser Leu His	
245	250	255
Ser Gly Cys Ala Thr Met Glu Ala Leu	Phe Phe Cys Tyr Cys Phe	
260	265	270
Pro Lys Leu Lys Pro Leu Phe Ile Ala	Tyr Val Cys Trp Leu Trp	
275	280	285
Trp Ser Thr Met Tyr Leu Thr His His	Tyr Phe Val Asp Leu Met	
290	295	300
Ala Gly Ser Val Leu Ser Tyr Val Ile	Phe Gln Tyr Thr Lys Tyr	
305	310	315
Thr His Leu Pro Ile Val Asp Thr Ser	Leu Phe Cys Arg Trp Ser	
320	325	330
Tyr Thr Ser Ile Glu Lys Tyr Asp Ile	Ser Lys Ser Asp Pro Leu	
335	340	345
Ala Ala Asp Ser Asn Asp Ile Glu Ser	Val Pro Leu Ser Asn Leu	
350	355	360
Glu Leu Asp Phe Asp Leu Asn Met Thr	Asp Glu Pro Ser Val Ser	
365	370	375
Pro Ser Leu Phe Asp Gly Ser Thr Ser	Val Ser Arg Ser Ser Ala	
380	385	390
Thr Ser Ile Thr Ser Leu Gly Val Lys	Arg Ala	
395	400	

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 422 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met	Ser	Ala	Leu	Ser	Thr	Leu	Lys	Lys	Arg	Leu	Ala	Ala	Cys	Asn
1				5					10					15
Arg	Ala	Ser	Gln	Tyr	Lys	Leu	Glu	Thr	Ser	Leu	Asn	Pro	Met	Pro
				20					25					30
Thr	Phe	Arg	Leu	Leu	Arg	Asn	Thr	Lys	Trp	Ser	Trp	Thr	His	Leu
				35					40					45
Gln	Tyr	Val	Phe	Leu	Ala	Gly	Asn	Leu	Ile	Phe	Ala	Cys	Ile	Val
				50					55					60
Ile	Glu	Ser	Pro	Gly	Phe	Trp	Gly	Lys	Phe	Gly	Ile	Ala	Cys	Leu
				65					70					75
Leu	Ala	Ile	Ala	Leu	Thr	Val	Pro	Leu	Thr	Arg	Gln	Ile	Phe	Phe
				80					85					90
Pro	Ala	Ile	Val	Ile	Ile	Thr	Trp	Ala	Ile	Leu	Phe	Tyr	Ser	Cys
				95					100					105
Arg	Phe	Ile	Pro	Glu	Arg	Trp	Arg	Pro	Pro	Ile	Trp	Val	Arg	Val
				110					115					120
Leu	Pro	Thr	Leu	Glu	Asn	Ile	Leu	Tyr	Gly	Ser	Asn	Leu	Ser	Ser
				125					130					135
Leu	Leu	Ser	Lys	Thr	Thr	His	Ser	Ile	Leu	Asp	Ile	Leu	Ala	Trp
				140					145					150
Val	Pro	Tyr	Gly	Val	Met	His	Tyr	Ser	Ala	Pro	Phe	Ile	Ile	Ser
				155					160					165

Phe Ile Leu Phe Ile Phe Ala Pro Pro Gly Thr Leu Pro Val Trp		
	170	175 180
Ala Arg Thr Phe Gly Tyr Met Asn Leu Phe Gly Val Leu Ile Gln		
	185	190 195
Met Ala Phe Pro Cys Ser Pro Pro Trp Tyr Glu Asn Met Tyr Gly		
	200	205 210
Leu Glu Pro Ala Thr Tyr Ala Val Arg Gly Ser Pro Gly Gly Leu		
	215	220 225
Ala Arg Ile Asp Ala Leu Phe Gly Thr Ser Ile Tyr Thr Asp Gly		
	230	235 240
Phe Ser Asn Ser Pro Val Val Phe Gly Ala Phe Pro Ser Leu His		
	245	250 255
Ala Gly Trp Ala Met Leu Glu Ala Leu Phe Leu Ser His Val Phe		
	260	265 270
Pro Arg Tyr Arg Phe Cys Phe Tyr Gly Tyr Val Leu Trp Leu Cys		
	275	280 285
Trp Cys Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Val		
	290	295 300
Gly Gly Met Cys Leu Ala Ile Ile Cys Phe Val Phe Ala Gln Lys		
	305	310 315
Leu Arg Leu Pro Gln Leu Gln Thr Gly Lys Ile Leu Arg Trp Glu		
	320	325 330
Tyr Glu Phe Val Ile His Gly His Gly Leu Ser Glu Lys Thr Ser		
	335	340 345
Asn Ser Leu Ala Arg Thr Gly Ser Pro Tyr Leu Leu Gly Arg Asp		
	350	355 360
Ser Phe Thr Gln Asn Pro Asn Ala Val Ala Phe Met Ser Gly Leu		
	365	370 375
Asn Asn Met Glu Leu Ala Asn Thr Asp His Glu Trp Ser Val Gly		
	380	385 390
Ser Ser Ser Pro Glu Pro Leu Pro Ser Pro Ala Ala Asp Leu Ile		
	395	400 405



Asp Arg Pro Ala Ser Thr Thr Ser Ser Ile Phe Asp Ala Ser His  
410 415 420  
Leu Pro

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 471 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Ala Ser Ser Ile Leu Arg Ser Lys Ile Ile Gln Lys Pro Tyr  
1 5 10 15  
Gln Leu Phe His Tyr Tyr Phe Leu Ser Glu Lys Ala Pro Gly Ser  
20 25 30  
Thr Val Ser Asp Leu Asn Phe Asp Thr Asn Ile Gln Thr Ser Leu  
35 40 45  
Arg Lys Leu Lys His His His Trp Thr Val Gly Glu Ile Phe His  
50 55 60  
Tyr Gly Phe Leu Val Ser Ile Leu Phe Phe Val Phe Val Val Phe  
65 70 75  
Pro Ala Ser Phe Phe Ile Lys Leu Pro Ile Ile Leu Ala Phe Ala  
80 85 90  
Thr Cys Phe Leu Ile Pro Leu Thr Ser Gln Phe Phe Leu Pro Ala  
95 100 105  
Leu Pro Val Phe Thr Trp Leu Ala Leu Tyr Phe Thr Cys Ala Lys  
110 115 120  
Ile Pro Gln Glu Trp Lys Pro Ala Ile Thr Val Lys Val Leu Pro  
125 130 135  
Ala Met Glu Thr Ile Leu Tyr Gly Asp Asn Leu Ser Asn Val Leu  
17

140	145	150
Ala Thr Ile Thr Thr Gly Val Leu Asp Ile Leu Ala Trp Leu Pro		
155	160	165
Tyr Gly Ile Ile His Phe Ser Phe Pro Phe Val Leu Ala Ala Ile		
170	175	180
Ile Phe Leu Phe Gly Pro Pro Thr Ala Leu Arg Ser Phe Gly Phe		
185	190	195
Ala Phe Gly Tyr Met Asn Leu Leu Gly Val Leu Ile Gln Met Ala		
200	205	210
Phe Pro Ala Ala Pro Pro Trp Tyr Lys Asn Leu His Gly Leu Glu		
215	220	225
Pro Ala Asn Tyr Ser Met His Gly Ser Pro Gly Gly Leu Gly Arg		
230	235	240
Ile Asp Lys Leu Leu Gly Val Asp Met Tyr Thr Thr Gly Phe Ser		
245	250	255
Asn Ser Ser Ile Ile Phe Gly Ala Phe Pro Ser Leu His Ser Gly		
260	265	270
Cys Cys Ile Met Glu Val Leu Phe Leu Cys Trp Leu Phe Pro Arg		
275	280	285
Phe Lys Phe Val Trp Val Thr Tyr Ala Ser Trp Leu Trp Trp Ser		
290	295	300
Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Ile Gly Gly		
305	310	315
Ala Met Leu Ser Leu Thr Val Phe Glu Phe Thr Lys Tyr Lys Tyr		
320	325	330
Leu Pro Lys Asn Lys Glu Gly Leu Phe Cys Arg Trp Ser Tyr Thr		
335	340	345
Glu Ile Glu Lys Ile Asp Ile Gln Glu Ile Asp Pro Leu Ser Tyr		
350	355	360
Asn Tyr Ile Pro Val Asn Ser Asn Asp Asn Glu Ser Arg Leu Tyr		
365	370	375
Thr Arg Val Tyr Gln Glu Ser Gln Val Ser Pro Pro Gln Arg Ala		

	380		385		390
Glu Thr Pro Glu Ala Phe Glu Met Ser Asn Phe Ser Arg Ser Arg					
	395		400		405
Gln Ser Ser Lys Thr Gln Val Pro Leu Ser Asn Leu Thr Asn Asn					
	410		415		420
Asp Gln Val Ser Gly Ile Asn Glu Glu Asp Glu Glu Glu Glu Gly					
	425		430		435
Asp Glu Ile Ser Ser Ser Thr Pro Ser Val Phe Glu Asp Glu Pro					
	440		445		450
Gln Gly Ser Thr Tyr Ala Ala Ser Ser Ala Thr Ser Val Asp Asp					
	455		460		465
Leu Asp Ser Lys Arg Asn					
	470				

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY:
- (B) LOCATION:
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION: /note= "Xaa at position 3 is Val or Ile"

(ix) FEATURE:

- (A) NAME/KEY:
- (B) LOCATION:
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION: /note= "Xaa at position 7 is Leu or Val"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Leu Asp Xaa Leu Ala Trp Xaa Pro Tyr Gly

1 5 10

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Phe Gly Ala Phe Pro Ser Leu His

1 5

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY:

(B) LOCATION:

(C) IDENTIFICATION METHOD:

(D) OTHER INFORMATION: /note= "Xaa at position 5 is Ser or Thr"

(ix) FEATURE:

(A) NAME/KEY:

(B) LOCATION:

(C) IDENTIFICATION METHOD:

(D) OTHER INFORMATION: /note= "Xaa at position 9 is Ala or Phe"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Thr Met Tyr Leu Xaa His His Tyr Xaa Val Asp Leu

1

5

10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2935 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
atattcttcc ccataacaac tcttctcgcc cttcctccgg ctccgtggcc aaattgtttt 60
atgcagcgcc tcctagcgat ttaacctcgt tctcgttgcc cttgcctgtc cgccttgcggt 120
cagtagcacc cttgcaacgt gaccttcccc agagtatcct cgtttgggcg ctggagaccg 180
gagcttgcac cctcataaac tagctcttcg aaatcaattc tccgttctcc agagattatc 240
ggatcgaatc tctccgctgt cgacaccttt cgtctctcgg tgatcctcgc ccttgaggatc 300
tcgtcacggt gacgccttga acccctggcc gccaaactcca cataggagac cacacttcat 360
tcttcccccg ccataattgc agcaccctcc gtctcccttc gagtcctcc tggatcatca 420
agtccgaaag gattagactc gtcgcagcga tgaataccac cttccatcc tggaaggatc 480
ggacgcaaaa ccagttcggc aagctccaga tccaagtccc atggcgcacc atacaacttc 540
tcgtgccgca ccgtatgcga cggaagattc ggtccaagct gcgcagtcgg atctcgccta 600
cctcatcgat atcctcgttg cagacgtcat tctcacctgt cgatacactc aggtcgctgc 660
aaagtcatag atggacgctc tatgactttc agtatctttt gctgctgatt gtcggcatat 720
tctcgctgag cgttatggaa tcacctggac cattggcaaa gaccgccgcg ttacgctac 780
ttctcgtctc tctccttctc ccgattacgc gccagttctt cttgccattc ctcccgattg 840
caggatggct tatatttttc tacgcttgcc agttcatccc gagcgactgg cgccctgcaa 900
tctgggttcg cgtgctgccg gctctggaaa acattctcta cggtgctaata atcagtaaca 960
tcctttccgc tcacaaaaat gtggtgcttg acgttttggc gtggcttccc tacggaatct 1020
```

gccattatgg cgcgccattt gtgtgctcag cgatcatgtt catcttttggc cctcccggca 1080  
ccgtccccct tttcgctcga acttttggat acatcagcat ggctgcagtc accattcagc 1140  
tggttttccc ctgctctcct ccgtggtacg aaaatctgta tggtttggct ccggctgatt 1200  
actccatgcc gggtaatcct gcgggccttg ctgcgcatcga tgagcttttt gggatagact 1260  
tgtacacatc gggcttcaga caatctcccg tcgtgttttg cgcatttcct tccctacatg 1320  
ccgtgattc gacacttgca gctctattta tgagccaagt gttcccacgg ttgaagccct 1380  
tgtttgtcat ctatactctc tggatgtggt gggctacaat gtatctttcg caccactacg 1440  
ctgttgatct ggtcggtggt ggcctcttgg caactgtcgc gttctacttt gctaaaacgc 1500  
ggttcatgcc tcgctccag aatgataaga tgttccgctg ggactacgat tatgttgagt 1560  
acggcgattc cgcactcgac tatgggtacg gtccagccag cttcgaaggc gaattcaacc 1620  
ttgatagcga tgagtggacc gttggttctt cgtcatccat ttcgtccggc tccctcagtc 1680  
cagttgacga ccactactct tgggagggcg agactcttgc ctctcctgcc accgacatcg 1740  
agtctggaag gcatttctga tcctgctcaa tgagccttga tacgtactac actgtgtacg 1800  
tgctactgca ttgactaatg agacggcggt ttcaaacaaa ttttaacgac atgcttggtt 1860  
atcgcattga gctgatttcg acacatatat atgtttaata cgttttgggg aactccagg 1920  
gattcatgac ggttgcttca tatcccgacc tggggatgga ttgacctggt tgtgcccaat 1980  
tttcttctgc ctaacgtttt gattatacat gcatttttca cgaaaccagc cggcccgcga 2040  
tgatcgtgac ctcaatttga gctcgaatct tcctggggcc tccagcgata attcttaatg 2100  
ctcgttccga ggggtgccaca tcggacattc gcttgtaaca cttttgcaga acgaacattt 2160  
tcaccgattc caacttgagt cattcgctta ctactttcaa ctggtcgaga aacttcgctt 2220  
cttttcagct cggctaggtg cataaatatt ttacattcgt gtcgatcgct cacatttcac 2280  
ggcgcctgga aacttggggg tttcgatttc attggaagg ataaacaaca tgggctgggc 2340  
gccttttaca cgcactacat tcgcttagaa aagttctgat gcttttaatg attcttgcac 2400  
tggcatatag aaaggggtcc tccagactcg ctactgtggt cctctctcaa cccacactc 2460  
gcttgctttt aacagtggac acccgtgga gctacgtctc catcaaataa ttggcatcaa 2520  
ccggaatcga tgccaggagg actgagctta ctacgggtga ccgtcgggta aaaggcgctt 2580  
atacagaaca tctcctctat cctcctgtcc catctcgatt ctctggctgc tggacgcaac 2640  
acacctgct gcacgttttc gacttcctaa tacgacctaa tatcattctc gggtttcttt 2700  
gctctggctc gcggccgcca tttatatggc gtgtcggctc gagtctgaag tgaacttttt 2760  
tctcctttct ggctccaca actttccgat ccctagcagc ttcccgtgca cagcgagggtg 2820  
ttgttggatg attgttccat agcattatca ttattcctaa tccggtagcg ttatgattta 2880  
tgaagaacag tgatgtacat tattatgcgg tgatcaaaaa aaaaaaaaaa aaaaa 2935

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2856 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iv) ANTI-SENSE: Yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
tagtatacaa gttgccttca taaatcataa cactaccgga ttataactaa taatactgct 60
gctatggata atttgcccca accacacaca cacacacact atccaaggga aacgaaaaga 120
agaggatagt gtccatgagc ggaatttgtg aggtgaggag gatgtgatag gtaggtgtga 180
gagcgtcaga ccatagcaag caagtcaagt caccgtccaa caagtcgaga atctgaggtt 240
agtcgaagcc ggggtcaaate gggacgcgtg aacatccttc aacaagcccc gtcagcgtca 300
gactctcctc ccaaccttgc acgtccggcg aatgtcagaa agcaacacag tagccatgtg 360
tatagcccc gggtaatatg tggacgtaat ccactttgag cgaatggata aaccatgcga 420
gcaatcaaga agaaggtcca tccttccctt ctttctatat gctaatgcag agttcgtgat 480
gcattgcaaa gcagaaatgt aaaaccgtct atgcacagcg agtgaagtgc aagaagaaaa 540
ggcgccccat tccgtgattt tttccaatat aaatgcttcc agtcccaagc gccaagagta 600
tgagcgatcg acacgaatgg taaaaataaa tatatagcat agtcgaaacg aataataagg 660
atgggttcga ccaggtcaag gtaatagcgg acgatagaag tgagaattga agtagtgtct 720
gggtggacac gagacaagcg agtgttaccg acgggggtta tcggatggtc ttccaggact 780
aaacaccagg atgcgtagcc ccattgtgtc cattaacatt atgatggagc aacggcgtgt 840
agcgatgagc gataacccat cccacgttcg agcccaaccc ggtttcgcaa tgtcaagcaa 900
tgtacagtga agtgcgggcg aaagagaaac agtatcccag gcggattcgg accgaggtac 960
attaaggctc aaccgatgat ctgggacacc gaagaaaggg tattaagat atatgtgtcg 1020
aaatgtgctc gatgcaataa agccagaatg tcgtgaaatg ttctaacgcc gtctcgtaaa 1080
tagatgcagt ggaggaggag cgagatcgct atatgcaggt atcgagtttg gtttgtggct 1140
actcaagggc tgaaagtatg cctgccggac tcaatatcag tatgtgggga ggtcagtgcc 1200
tcggtttccc atgagtaatg atcgtcaacg ggactcaagg agcctgagga gacggaggat 1260
```

```

gaagaaccaa cagtccattc atcgtgtcg agattgaagt ctccatcata gccagctgca 1320
ccatacccat actccagggc agactcgccg aattccacat agtcgtagtc ccaacggaag 1380
gtcttgtcga gctggacacg gggaaggaat cgggtcttgg cgaagtagaa agcaatggcg 1440
gccaggagac caccgcgaac caaatcgacc gcatagtgat gtgagaggta cattgttgcc 1500
caccacatcc atagagtata ggtcacgaag acgggcttca tgcgggggaa aacatgactc 1560
atgaaaagtg cggccagggg tgagtcggca gcatgcagcg acggaaaagc gccgaacaca 1620
acaggcgact gatggaaaac agacgtgtaa aggtcgatgc cgaaaagctt gtcaatgcgg 1680
gcaagccctg cgggatcacc ttggatggag tagtctgccg gagctagacc atagcgattc 1740
tcataccaag gtggagagca agggaaaaac agctgaatag taaccgcagt catactgata 1800
tagccgaaag tgcgcgcgaa aaggggaaca gtgccgggcg gaccgaagat gaacatgata 1860
aacgagcaca caaacggagc gccatagtgg cagataccgt agggtagcca cgccagcacg 1920
tcaagcacia cgttctggtg agcggatagg atgttgctga tgtttgccg gtagagaata 1980
ttctccagt caggcaagac acgaaccaa atcgcagggc gccaatcgct tgggatgaac 2040
tggcaggcgt agaaaaacag aagccatccg gcaatcggca gaaacgggag gaagaactgg 2100
cgggtcatag ggatcaggag agagaatagg agcatggaga aaatggccgt tttgccccaa 2160
ggcccgggcg actcgataac ggtcaaagag aagatgccca cgatcaacag aagcagatat 2220
tggaagtcgt aaaccgtcca tcggtggctt tggagcgatc gtagtgtgtc tgcaggcgat 2280
aacgacgtct gtaaagaggc tattgacgag gtaggagacg ctctactgcg caatttggac 2340
cttaacttcc gccgcatgcg atgcgggacg agcagttgga tggaccgcca tggaacctgg 2400
atctgaagct ttccaaactg gttctgctg cggtccttcc acgtgggaag tgtttgattc 2460
atggttgccg cgagtcttgc gggcaagact cagctataat tcctccagat cgaactccgt 2520
ctgctttctt tcccccgat tgtacaagag ttggaaggcg acgaataaat cggtcgcgga 2580
agcggcagaa aaggggtctt gaataacaag ccgttaggca acgagcaggg tgctgacggc 2640
gtcaaccgga gcggggaaat tgagatcggg tcgaagaaga cgaaacgaag cgcacgctcc 2700
ggtctccagc ggccaaacga ggatgtctt gagccgtcac gttgcagggg gtttactga 2760
tgcaaggcga agtgcaaggc cgcagcagtc gagaggaaaa cctggaacga ggaggtcgtg 2820
agggaggcag atggccacgg agccggagta taaacc 2856

```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2856 bases

(B) TYPE: nucleic acid



- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: mRNA
- (iv) ANTI-SENSE: Yes
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

uaguauacaa guugccuuca uaaaucauaa cacuaccgga uuauaacuaa uaauacugcu   60
gcuauaggaua auuugcccca accacacaca cacacacacu auccaaggga aacgaaaaga  120
agaggauagu guccaugagc ggaauuugug aggugaggag gaugugauag guagguguga  180
gagcgucaga ccuagcaag caagucaagu caccguccaa caagucgaga aucugagguu  240
agucgaagcc gggucaaauc gggacgcgug aacauccuuc aacaagcccc gucagcguca  300
gacucuccuc ccaaccuugc acguccggcg aaugucagaa agcaacacag uagccaugug  360
uauagcccc ggguaauaug uggacguaau ccacuuugag cgaauaggaua aaccaugcga  420
gcaaucaaga agaaggucca uccuuccccu cuuucuaauu gcuaaugcag aguucgugau  480
gcauugcaaa gcagaaaugu aaaaccgucu augcacagcg agugaagugc aagaagaaaa  540
ggcgccccau uccgugauuu uuuccaauu aaaugcuucc agucccaagc gccaagagua  600
ugagcgaucg acacgaauug uaaaaauaaa uauauagcau agucgaaacg aauaauaagg  660
auggguuuga ccaggucaag guaauagcgg acgauagaag ugagaauuga aguagugucu  720
ggguggacac gagacaagcg aguguuaccg acggggguua ucggaugguc uuccaggacu  780
aaacaccagg augcguagcc ccauuguguc cauuacauu augauggagc aacggcgugu  840
agcgaugagc gauaaccuau cccacguucg agcccaaccc gguuucgcaa ugucaagcaa  900
uguacaguga agugcggcg aaagagaaac aguauccag gcggaauucg accgagguac  960
auuaaggucg aaccgaugau cugggacacc gaagaaaggg uauuaaagau auaugugucg 1020
aaugugcuc gaugcaauaa agccagaauug ucgugaaaug uucuaacgcc gucucguaaa 1080
uagaugcagu ggaggaggag cgagaucgcu auaugcaggu aucgaguug guuuguggcu 1140
acucaagggc ugaaaguau ccugccggac ucaauaucag uaugugggga ggucagugcc 1200
ucgguuuccc augaguauug aucgucaacg ggacucaagg agccugagga gacggaggau 1260
gaagaaccaa caguccauuc aucgcugucg agauugaagu cuccaucua gccagcugca 1320
ccauaccuau acuccagggc agacucgccg aaauccacau agucguaguc ccaacggaag 1380
gucuugucga gcuggacacg gggaaaggaau cgggucuugg cgaaguagaa agcaauggcg 1440
gccaggagac caccgcgaac caaauccgacc gcauagugau gugagaggua cauuguugcc 1500
caccacaucc auagaguaua ggucacgaag acgggcuuca ugcgggggaa acaugacuc 1560

```

```

augaaaagug cggccagggg uagagucggca gcaugcagcg acggaagc gccgaacaca 1620
acaggcgacu gauggaaaac agacguguaa aggucgaugc cgaaaagcuu gucaaugcgg 1680
gcaagcccug cgggaucacc uuggauggag uagucugccg gagcuagacc auagcgauuc 1740
ucauaccaag guggagagca agggaaaaac agcugaauag uaaccgcagu cauacugaua 1800
uagccgaaag ugcgcgcgaa aaggggaaca gugccgggcg gaccgaagau gaacaugauc 1860
aacgagcaca caaacggagc gccauagugg cagauaccgu aggguaagcca cgccagcacg 1920
ucaagcaca cguucuggug agcggauagg auguugcuga uguuugcgcc guagagaaua 1980
uucuccagug caggcaagac acgaacccaa aucgcagggc gccaaucgcu ugggaugaac 2040
uggcaggcgu agaaaaacag aagccauccg gcaaucggca gaaacgggag gaagaacugg 2100
cgggucauag ggaucaggag agagaauagg agcauggaga aaauggccgu uuugcccaa 2160
ggcccgggcg acucgauaac ggucaaagag aagaugccca cgaucaacag aagcagauau 2220
uggaagucgu aaaccgucca ucgguggcuu uggagcgau guaguguguc ugcaggcgau 2280
aacgacgucu guaaagaggc uauugacgag guaggagacg cucuacugcg caauuuggac 2340
cuuaacuucc gccgcaugcg augcgggacg agcaguugga uggaccgcca uggaaccugg 2400
aucugaagcu uuccaaacug guucugcgug cgguccuucc acgugggaag uguuugauuc 2460
augguugcgg cgagucuugc gggcaagacu cagcuauaa uccuccagau cgaacuccgu 2520
cugcuuucuu ucccccgau uguacaagag uuggaaggcg acgaauaaau cggucgcgga 2580
agcggcagaa aaggggucuu gaauaacaag ccguuaggca acgagcaggg ugcugacggc 2640
gucaaccga gcggggaaau ugagaucggu ucgaagaaga cgaaacgaag cgcacgcucc 2700
ggucuccagc ggccaaacga ggaugucuuu gagccgucac guugcagggg guuucacuga 2760
ugcaaggcga agugcaaggg cgcagcaguc gagaggaaaa ccuggaacga ggaggucgug 2820
agggaggcag auggccacgg agccggagua uaaacc 2856

```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2385 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

aagctttttt	gcctctgcaa	aagttccttt	ctcgaattgg	ttttttgagg	aaaagcaagt	60
taataaacta	attatattat	atataattag	caattttata	aaaaaaataa	aaaaatagcc	120
ctgattgctg	gcaactgtga	gctgaacatt	ggttaatcgg	tccatctttt	tttaaataatt	180
ttacatcgct	acttttaagt	gcttgacact	tgcatttaat	agctactttc	tttccttcat	240
aaaaattcct	tttttttcct	ttagttttcc	ggttaattcc	ttacgaaatt	tttttcgtac	300
gcttcccttt	tttactctga	taattctttg	aagcaatgtc	tgctctttcg	accttaaaaa	360
agcgccttgc	tgcggtgaac	cgagcatccc	aatacaagtt	ggaaacaagc	ttaaacccta	420
tgccctacatt	tcgtttgcta	cgcaatacga	aatggtcattg	gacacatttg	caatatgtgt	480
ttctagcagg	taatttgatt	tttgcttgta	ttgtcattga	atctcctgga	ttctggggga	540
aatttggcat	tgccctgtctt	ttggccattg	cgttgaccgt	tcctttaaca	cgccaaattt	600
tttttcctgc	cattgttatc	atcacctggg	caattttatt	ttactcttgt	aggtttattc	660
cagaacgctg	gcgtccaccc	atatgggttc	gtgttttacc	cacacttgaa	aatattcttt	720
atggctctaa	tctttctagt	cttctctcga	aaaccacgca	tagcatcctt	gatatttttg	780
cctgggttcc	atatggagtc	atgcattatt	cggctccttt	tatcatttca	tttattcttt	840
tcattctttg	acctcctgga	actcttccag	tttgggctcg	aacttttggt	tatatgaatt	900
tatttgggtg	tcttatccaa	atggctttcc	cctgttctcc	tccttggtat	gaaaatatgt	960
atggtttaga	acctgccacg	tatgcagtac	gtggctctcc	tggtggattg	gcccgtattg	1020
atgctctctt	cggcactagc	atttactactg	attgtttttc	taactctccg	gttggtttttg	1080
gtgcctttcc	atctcttcac	gctggatggg	ccatgctgga	agcacttttc	ctttcgcattg	1140
tgtttcctcg	ataccgcttc	tgcttttatg	gatattgttct	atggctttgc	tggtgtacta	1200
tgtaccttac	ccaccactac	tttgtagatt	tggtcggcgg	tatgtgttta	gctattatat	1260
gcttcgtttt	tgctcaaaaag	ctacgcctcc	cacagttgca	aactggtaaa	atccttcggt	1320
gggaatacga	gtttgttatc	cacggtcattg	gtctttccga	aaaaaccagc	aactccttgg	1380
ctcgtaccgg	cagcccatac	ttacttggaa	gggattcttt	tactcaaaac	cctaattgcag	1440
tagccttcat	gagtgggtctt	aacaatatgg	aacttgctaa	caccgatcat	gaatgggtccg	1500
tggttctcatc	atcacctgag	ccgttaccta	gtcctgctgc	tgatttgatt	gatcgtcctg	1560
ccagtaccac	ttcctccatc	tttgatgcaa	gtcatcttcc	ttaaatcaac	gtgctttaag	1620
aatatatttc	caaaagctac	atgatacatt	gactagaatc	ggtttgattc	atagtgggtat	1680
tggaatgatg	ttgttcattg	tgttttttaa	ctgttaatct	gacatccatt	gagtcattct	1740
ttacaatttg	taaaattaat	ttgtatcact	aattttgaag	gaagctattt	tggtattaat	1800
accgcttttg	gtctccactt	ccttttcgaa	actcttaaca	gcgattaggc	cgggtatctt	1860
ccagtgtgat	gtataggtat	ttgtcgtttt	tttatcattt	ccgttaataa	agaactcttt	1920

```

tatccagctt cttacactgt caactgttgt gaaaggaaca catttagaat ttcattttcc 1980
ttatttggtg tgatttaaatt cgtttgacat aattttaaat ttggtttgaa atgtgtgtga 2040
gaaggcttgt tttattcatt tagttttattg cttgtttgca cgaaaatcca gaacggagca 2100
ttaatgtaat ccttttttat tctgtaaagc gtttttatac aaatgttggt tatacgtttc 2160
taaaataaga atattgttat aataatatag ttttttctat catttgttac acacactaaa 2220
gagacattaa ggataagcaa atgtgttaaa atgataatat attttggaag catttataaa 2280
gaaattaagc agctttgact aactacattt ttgttttttt cctaagcaaa actgtatagt 2340
tatacacgcg agctgtattc acttccattg tagtgacttg agctc 2385

```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 422 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```

Met Ser Ala Leu Ser Thr Leu Lys Lys Arg Leu Ala Ala Cys Asn
  1             5             10             15
Arg Ala Ser Gln Tyr Lys Leu Glu Thr Ser Leu Asn Pro Met Pro
             20             25             30
Thr Phe Arg Leu Leu Arg Asn Thr Lys Trp Ser Trp Thr His Leu
             35             40             45
Gln Tyr Val Phe Leu Ala Gly Asn Leu Ile Phe Ala Cys Ile Val
             50             55             60
Ile Glu Ser Pro Gly Phe Trp Gly Lys Phe Gly Ile Ala Cys Leu
             65             70             75
Leu Ala Ile Ala Leu Thr Val Pro Leu Thr Arg Gln Ile Phe Phe
             80             85             90
Pro Ala Ile Val Ile Ile Thr Trp Ala Ile Leu Phe Tyr Ser Cys
             95             100             105

```

Arg Phe Ile Pro Glu Arg Trp Arg Pro Pro Ile Trp Val Arg Val		
	110	115 120
Leu Pro Thr Leu Glu Asn Ile Leu Tyr Gly Ser Asn Leu Ser Ser		
	125	130 135
Leu Leu Ser Lys Thr Thr His Ser Ile Leu Asp Ile Leu Ala Trp		
	140	145 150
Val Pro Tyr Gly Val Met His Tyr Ser Ala Pro Phe Ile Ile Ser		
	155	160 165
Phe Ile Leu Phe Ile Phe Ala Pro Pro Gly Thr Leu Pro Val Trp		
	170	175 180
Ala Arg Thr Phe Gly Tyr Met Asn Leu Phe Gly Val Leu Ile Gln		
	185	190 195
Met Ala Phe Pro Cys Ser Pro Pro Trp Tyr Glu Asn Met Tyr Gly		
	200	205 210
Leu Glu Pro Ala Thr Tyr Ala Val Arg Gly Ser Pro Gly Gly Leu		
	215	220 225
Ala Arg Ile Asp Ala Leu Phe Gly Thr Ser Ile Tyr Thr Asp Cys		
	230	235 240
Phe Ser Asn Ser Pro Val Val Phe Gly Ala Phe Pro Ser Leu His		
	245	250 255
Ala Gly Trp Ala Met Leu Glu Ala Leu Phe Leu Ser His Val Phe		
	260	265 270
Pro Arg Tyr Arg Phe Cys Phe Tyr Gly Tyr Val Leu Trp Leu Cys		
	275	280 285
Trp Cys Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Val		
	290	295 300
Gly Gly Met Cys Leu Ala Ile Ile Cys Phe Val Phe Ala Gln Lys		
	305	310 315
Leu Arg Leu Pro Gln Leu Gln Thr Gly Lys Ile Leu Arg Trp Glu		
	320	325 330
Tyr Glu Phe Val Ile His Gly His Gly Leu Ser Glu Lys Thr Ser		
	335	340 345

Asn Ser Leu Ala Arg Thr Gly Ser Pro Tyr Leu Leu Gly Arg Asp		
	350	355 360
Ser Phe Thr Gln Asn Pro Asn Ala Val Ala Phe Met Ser Gly Leu		
	365	370 375
Asn Asn Met Glu Leu Ala Asn Thr Asp His Glu Trp Ser Val Gly		
	380	385 390
Ser Ser Ser Pro Glu Pro Leu Pro Ser Pro Ala Ala Asp Leu Ile		
	395	400 405
Asp Arg Pro Ala Ser Thr Thr Ser Ser Ile Phe Asp Ala Ser His		
	410	415 420
Leu Pro		

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2385 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```

aagctttttt gcctctgcaa aagttccttt ctcgaattgg ttttttgagg aaaagcaagt 60
taataaacta attatattat atataattag caattttata aaaaaataa aaaaatagcc 120
ctgattgctg gcaactgtga gctgaacatt ggtaaatcgg tccatctttt tttaaattatt 180
ttacatcgct acttttaagt gcttgacact tgcatttaat agctactttc tttccttcat 240
aaaaattcct tttttttcct ttagttttcc ggtaattcc ttacgaaatt tttttcgtag 300
gcttcccttt tttactctga taattctttg aagcaatgtc tgctctttcg accttaaaaa 360
agcgccttgc tgcgtgtaac cgagcatccc aatacaagtt ggaaacaagc ttaaacccta 420
tgcctacatt tcgtttgcta cgcaatacga aatgggtcatg gacacatttg caatatgtgt 480
ttctagcagg taatttgatt tttgcttgta ttgtcattga atctcctgga ttctggggga 540
aatttggcat tgcctgtctt ttggccattg cgttgaccgt tcctttaaca cgccaaattt 600
ttttcctgc cattgttatc atcacctggg caattttatt ttactcttgt aggtttattc 660

```

```

cagaacgctg gcgtccaccc atatgggttc gtgttttacc cacacttgaa aatattcttt 720
atggctctaa tctttctagt cttctctcga aaaccacgca tagcatcctt gatattttgg 780
cctgggttcc atatggagtc atgcattatt cggctccttt tatcatttca tttattcttt 840
tcatctttgc acctcctgga actcttccag tttgggctcg aacttttggt tatatgaatt 900
tatttggtgt tcttatccaa atggctttcc cctgttctcc tccttggtat gaaaatatgt 960
atggtttaga acctgccacg tatgcagtac gtggctctcc tgggtggattg gcccgatttg 1020
atgctctctt cggcactagc atttacctg atgggttttc taactctccg gttgtttttg 1080
gtgcctttcc atctcttcac gctggatggg ccatgctgga agcacttttc ctttcgcatg 1140
tgtttcctcg ataccgcttc tgcttttatg gatatgttct atggctttgc tgggtgtacta 1200
tgtaccttac ccaccactac tttgtagatt tggtcggcgg tatgtgttta gctattatat 1260
gcttcgtttt tgctcaaaag ctacgcctcc cacagttgca aactggtaaa atccttcgtt 1320
gggaatacga gtttgttatc cacggtcattg gtctttccga aaaaaccagc aactccttgg 1380
ctcgtaccgg cagcccatat ttacttgaa gggattcttt tactcaaaac cctaatgcag 1440
tagccttcat gagtggctct aacaatatgg aacttgctaa caccgatcat gaatggtccg 1500
tgggttcatc atcacctgag ccgttaccta gtcctgctgc tgatttgatt gatcgtcctg 1560
ccagtaccac ttctccatc tttgatgcaa gtcattctcc ttaaataaac gtgctttaag 1620
aatatatttc caaaagctac atgatacatt gactagaatc ggtttgattc atagtgggtat 1680
tggaatgatg ttgttcattg tgttttttaa ctgttaatct gacatccatt gagtcattct 1740
ttacaatttg taaaattaat ttgtatcact aattttgaag gaagctattt tgggtattaat 1800
accgcttttg gtctccactt ctttttcgaa actcttaaca gcgattaggc cgggtatctt 1860
ccagtgtgat gtataggtat ttgtcgtttt tttatcattt ccgttaataa agaactcttt 1920
tatccagctt cttacctgt caactgttgt gaaaggaaca catttagaat ttcattttcc 1980
ttatttggtg tgatttaaat cgtttgacat aattttaaat ttggtttgaa atgtgtgtga 2040
gaaggcttgt tttattcatt tagtttattg cttgtttgca cgaaaatcca gaacggagca 2100
ttaatgtaat ctttttttat tctgtaaagc gtttttatac aaatgttggt tatacgtttc 2160
taaaataaga atattgttat aataatatag ttttttctat catttggtac acacactaaa 2220
gagacattaa ggataagcaa atgtgttaaa atgataatat attttgghaaa catttataaa 2280
gaaattaagc agctttgact aactacattt ttgttttttt cctaagcaaa actgtatagt 2340
tatacacgcg agctgtattc acttccattg tagtgacttg agctc 2385

```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met	Ser	Ala	Leu	Ser	Thr	Leu	Lys	Lys	Arg	Leu	Ala	Ala	Cys	Asn
1			5						10					15
Arg	Ala	Ser	Gln	Tyr	Lys	Leu	Glu	Thr	Ser	Leu	Asn	Pro	Met	Pro
			20						25					30
Thr	Phe	Arg	Leu	Leu	Arg	Asn	Thr	Lys	Trp	Ser	Trp	Thr	His	Leu
			35						40					45
Gln	Tyr	Val	Phe	Leu	Ala	Gly	Asn	Leu	Ile	Phe	Ala	Cys	Ile	Val
			50						55					60
Ile	Glu	Ser	Pro	Gly	Phe	Trp	Gly	Lys	Phe	Gly	Ile	Ala	Cys	Leu
			65						70					75
Leu	Ala	Ile	Ala	Leu	Thr	Val	Pro	Leu	Thr	Arg	Gln	Ile	Phe	Phe
			80						85					90
Pro	Ala	Ile	Val	Ile	Ile	Thr	Trp	Ala	Ile	Leu	Phe	Tyr	Ser	Cys
			95						100					105
Arg	Phe	Ile	Pro	Glu	Arg	Trp	Arg	Pro	Pro	Ile	Trp	Val	Arg	Val
			110						115					120
Leu	Pro	Thr	Leu	Glu	Asn	Ile	Leu	Tyr	Gly	Ser	Asn	Leu	Ser	Ser
			125						130					135
Leu	Leu	Ser	Lys	Thr	Thr	His	Ser	Ile	Leu	Asp	Ile	Leu	Ala	Trp
			140						145					150
Val	Pro	Tyr	Gly	Val	Met	His	Tyr	Ser	Ala	Pro	Phe	Ile	Ile	Ser
			155						160					165
Phe	Ile	Leu	Phe	Ile	Phe	Ala	Pro	Pro	Gly	Thr	Leu	Pro	Val	Trp
			170						175					180
Ala	Arg	Thr	Phe	Gly	Tyr	Met	Asn	Leu	Phe	Gly	Val	Leu	Ile	Gln



	185	190	195
Met Ala Phe Pro Cys Ser Pro Pro Trp Tyr Glu Asn Met Tyr Gly			
	200	205	210
Leu Glu Pro Ala Thr Tyr Ala Val Arg Gly Ser Pro Gly Gly Leu			
	215	220	225
Ala Arg Ile Asp Ala Leu Phe Gly Thr Ser Ile Tyr Thr Asp Gly			
	230	235	240
Phe Ser Asn Ser Pro Val Val Phe Gly Ala Phe Pro Ser Leu His			
	245	250	255
Ala Gly Trp Ala Met Leu Glu Ala Leu Phe Leu Ser His Val Phe			
	260	265	270
Pro Arg Tyr Arg Phe Cys Phe Tyr Gly Tyr Val Leu Trp Leu Cys			
	275	280	285
Trp Cys Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Val			
	290	295	300
Gly Gly Met Cys Leu Ala Ile Ile Cys Phe Val Phe Ala Gln Lys			
	305	310	315
Leu Arg Leu Pro Gln Leu Gln Thr Gly Lys Ile Leu Arg Trp Glu			
	320	325	330
Tyr Glu Phe Val Ile His Gly His Gly Leu Ser Glu Lys Thr Ser			
	335	340	345
Asn Ser Leu Ala Arg Thr Gly Ser Pro Tyr Leu Leu Gly Arg Asp			
	350	355	360
Ser Phe Thr Gln Asn Pro Asn Ala Val Ala Phe Met Ser Gly Leu			
	365	370	375
Asn Asn Met Glu Leu Ala Asn Thr Asp His Glu Trp Ser Val Gly			
	380	385	390
Ser Ser Ser Pro Glu Pro Leu Pro Ser Pro Ala Ala Asp Leu Ile			
	395	400	405
Asp Arg Pro Ala Ser Thr Thr Ser Ser Ile Phe Asp Ala Ser His			
	410	415	420

Leu Pro

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2340 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
tttctttctg tcaaagaata ataaagtgcc catcagtgtt catatttggt acaaagtggg 60
tttctgattt ggtactactg cagaggcgta ttttttgctt cagttaccat agcgtaagaa 120
cactagcgac ttttgttcgt gaaccaacag agtaggattt ctactgctac atctcttagg 180
tagttggtta gtccgatcgc tcacttttgg ttgttggtta gtacttcata agtttatacct 240
tttccttttt cacactgagc tactttgggt atagcttttg gcccaaggat ctttgaattt 300
tctccaaaag tactttatatt tatatcctac aggttgcggt tttcatattt taaaaagctt 360
tttaatcatt cttttgcgta tggcaaacc tttttcgaga tggtttctat cagagagacc 420
tccaaactgc catgtagccg atttagaaac aagtttagat ccccatcaaa cgttggttgaa 480
ggtgcaaaaa tacaaaccg ctttaagcga ctgggtgcat tacatcttct tgggatccat 540
catgctgttt gtgttcatta ctaatccgc accttgatc ttcaagatcc ttttttattg 600
tttcttgggc actttattca tcattccagc tacgtcacag tttttcttca atgccttgcc 660
catcctaaca tgggtggcgc tgtatttcac ttcacgtac tttccagatg accgcaggcc 720
tcctattact gtcaaagtgt taccagcggg ggaacaatt ttatacggcg acaatttaag 780
tgatattctt gcaacatcga cgaattcctt tttggacatt ttagcatggg taccgtacgg 840
actatttcat tatggggccc catttgctgt tgctgccatc ttattcgtat ttgggtccacc 900
aactgttttg caagggtatg cttttgcatt tgggttatatg aacctgtttg gtgttatcat 960
gcaaaatgtc tttccagccg ctcccccag gtataaaatt ctctatggat tgcaatcagc 1020
caactatgat atgcatgggt cgcctgggtg attagctaga attgataagc tactcggtat 1080
taatatgtat actacagctt tttcaaattc ctccgtcatt ttcggtgctt ttccttcact 1140
gcattccggg tgtgctacta tggaagccct gtttttctgt tattgttttc caaaattgaa 1200
gcccttgttt attgcttatg tttgctggtt atgggtgggtca actatgtatc tgacacacca 1260
ttattttgta gaccttatgg caggttctgt gctgtcatatc gttattttcc agtacacaaa 1320
```

gtacacacat ttaccaattg tagatacatc tcttttttgc agatgggtcat acacttcaat 1380  
 tgagaaatac gatatatcaa agagtgatcc attggctgca gattcaaacg atatcgaaag 1440  
 tgtccctttg tccaacttgg aacttgactt tgatcttaat atgactgatg aaccagtggt 1500  
 aagcccttcg ttatttgatg gatctacttc tgtttctcgt tcgtccgcca cgtctataac 1560  
 gtcactaggt gtaaagaggg cttaatgagt attttatctg caattacgga tacggttgggt 1620  
 cttatgtaga tacatataaa tatatatctt tttctttctt tttcttagtc aggattgtcg 1680  
 ttttagcataa tatacatgta gtttatttaa tcacatacca ctgattatct ttagaatttt 1740  
 ataaattttt gaaataaatg ggtggctttt aatgggtgtct atgttaagtg aggccttttag 1800  
 aatgctcttc ctgctttggt tattatatgt gtatgaaaga tatgtatgta tttacatgtg 1860  
 tttgtagcgt cccagtcga aacctgtgcg ctatacctaa atggattgat aatcttcatt 1920  
 cactaattct aaaatagact tcttcccaa agaacgggtg aacgatgagg ctctatccag 1980  
 ctgcttatct aaatcaactt taacgatgga tgatcttatg acacggggat ctttctttaa 2040  
 agttcttaga atttcagact gtaccgcagc tgatgaatca aacagcatta aaaagtgata 2100  
 tgctcgaaaa tgtttttcct ggtctttctt cattatttta ggaagatacc ttatgccccat 2160  
 ggggtacaatg tccctcacca cacctctggt ttgaataatc agtttcccga ttgtggaaga 2220  
 caattctttt gcttccaact ttggcgcatt ggagttgggt atgcgaacaa gtccgatcag 2280  
 ctcataaagc atcttagtga aaagggtgggt tttgcgttat tctttcctct gttgaagctt 2340

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 401 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met	Ala	Asn	Pro	Phe	Ser	Arg	Trp	Phe	Leu	Ser	Glu	Arg	Pro	Pro
1				5				10					15	
Asn	Cys	His	Val	Ala	Asp	Leu	Glu	Thr	Ser	Leu	Asp	Pro	His	Gln
				20				25					30	
Thr	Leu	Leu	Lys	Val	Gln	Lys	Tyr	Lys	Pro	Ala	Leu	Ser	Asp	Trp
								35						

35	40	45
Val His Tyr Ile Phe Leu Gly Ser Ile Met Leu Phe Val Phe Ile		
50	55	60
Thr Asn Pro Ala Pro Trp Ile Phe Lys Ile Leu Phe Tyr Cys Phe		
65	70	75
Leu Gly Thr Leu Phe Ile Ile Pro Ala Thr Ser Gln Phe Phe Phe		
80	85	90
Asn Ala Leu Pro Ile Leu Thr Trp Val Ala Leu Tyr Phe Thr Ser		
95	100	105
Ser Tyr Phe Pro Asp Asp Arg Arg Pro Pro Ile Thr Val Lys Val		
110	115	120
Leu Pro Ala Val Glu Thr Ile Leu Tyr Gly Asp Asn Leu Ser Asp		
125	130	135
Ile Leu Ala Thr Ser Thr Asn Ser Phe Leu Asp Ile Leu Ala Trp		
140	145	150
Leu Pro Tyr Gly Leu Phe His Tyr Gly Ala Pro Phe Val Val Ala		
155	160	165
Ala Ile Leu Phe Val Phe Gly Pro Pro Thr Val Leu Gln Gly Tyr		
170	175	180
Ala Phe Ala Phe Gly Tyr Met Asn Leu Phe Gly Val Ile Met Gln		
185	190	195
Asn Val Phe Pro Ala Ala Pro Pro Trp Tyr Lys Ile Leu Tyr Gly		
200	205	210
Leu Gln Ser Ala Asn Tyr Asp Met His Gly Ser Pro Gly Gly Leu		
215	220	225
Ala Arg Ile Asp Lys Leu Leu Gly Ile Asn Met Tyr Thr Thr Ala		
230	235	240
Phe Ser Asn Ser Ser Val Ile Phe Gly Ala Phe Pro Ser Leu His		
245	250	255
Ser Gly Cys Ala Thr Met Glu Ala Leu Phe Phe Cys Tyr Cys Phe		
260	265	270
Pro Lys Leu Lys Pro Leu Phe Ile Ala Tyr Val Cys Trp Leu Trp		

	275	280	285
Trp Ser Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Met			
	290	295	300
Ala Gly Ser Val Leu Ser Tyr Val Ile Phe Gln Tyr Thr Lys Tyr			
	305	310	315
Thr His Leu Pro Ile Val Asp Thr Ser Leu Phe Cys Arg Trp Ser			
	320	325	330
Tyr Thr Ser Ile Glu Lys Tyr Asp Ile Ser Lys Ser Asp Pro Leu			
	335	340	345
Ala Ala Asp Ser Asn Asp Ile Glu Ser Val Pro Leu Ser Asn Leu			
	350	355	360
Glu Leu Asp Phe Asp Leu Asn Met Thr Asp Glu Pro Ser Val Ser			
	365	370	375
Pro Ser Leu Phe Asp Gly Ser Thr Ser Val Ser Arg Ser Ser Ala			
	380	385	390
Thr Ser Ile Thr Ser Leu Gly Val Lys Arg Ala			
	395	400	

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2340 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```

tttctttctg tcaaagaata ataaagtgcc catcagtgtt catatttggtt acaaagtggg 60
tttctgattt ggtactactg cagaggcgta ttttttgctt cagttaccat agcgtaagaa 120
cactagcgac ttttgttcgt gaaccaacag agtaggattt ctactgctac atctcttagg 180
tagttggtta gtccgatcgc tcacttttgg ttgttggttaa gtacttcata agtttatcct 240
tttccttttt cacactgagc tactttgggt atagcttttg gcccaaggat ctttgaattt 300

```

tctccaaaag tactttattht tatatcttac aggttgctgt tttcatattht taaaaagctt 360  
tttaatcatt cctttgctga tggcaaacc tttttcgaga tggtttctat cagagagacc 420  
tccaaactgc catgtagccg atttagaaac aagtttagat ccccatcaaa cgttggtgaa 480  
ggtgcaaaaa tacaaaccg ctttaagcga ctgggtgcat tacatcttct tgggatccat 540  
catgctgttt gtgttcatta ctaatccgc accttgatc ttcaagatcc ttttttattg 600  
tttcttgggc actttattca tcattccagc tacgtcacag tttttcttca atgccttgcc 660  
catcctaaca tgggtggcgc tgtatttcac ttcacgtac tttccagatg accgcaggcc 720  
tcctattact gtcaaagtgt taccagcgtt ggaaacaatt ttatacggcg acaatttaag 780  
tgatattctt gcaacatcga cgaattcctt tttggacatt ttagcatggt taccgtacgg 840  
actatttcat tttggggccc catttgctgt tgctgccatc ttattcgtat ttggtccacc 900  
aactgttttg caaggttatg cttttgcatt tggttatatg aacctgtttg gtgttatcat 960  
gcaaaatgtc tttccagccg ctcccccag gtataaaatt ctctatggat tgcaatcagc 1020  
caactatgat atgcatggct cgcctggtgg attagctaga attgataagc tactcggtat 1080  
taatattgat actacagctt tttcaaattc ctccgtcatt ttcggtgctt ttccttcact 1140  
gcattccggg tgtgctacta tggaagccct gtttttctgt tattgttttc caaaattgaa 1200  
gcccttgtht attgcttatg tttgctggtt atgggtggtca actatgtatc tgacacacca 1260  
ttattttgta gaccttatgg caggttctgt gctgtcatac gttattttcc agtacacaaa 1320  
gtacacacat ttaccaattg tagatacatc tcttttttgc agatgggtcat acacttcaat 1380  
tgagaaatac gatatatcaa agagtgatcc attggctgca gattcaaacg atatcgaaag 1440  
tgtccctttg tccaacttg aacttgactt tgatcttaat atgactgatg aaccagtggt 1500  
aagcccttcg ttatttgatg gatctacttc tgtttctcgt tcgtccgcca cgtctataac 1560  
gtcactaggt gtaaagagg cttaatgagt attttatctg caattacgga tacggttggt 1620  
cttatgtaga tacatataaa tatatatctt tttctttctt tttcttagtc aggattgtcg 1680  
tttagcataa tatacatgta gtttatthta tcacatacca ctgattatct ttagaatttt 1740  
ataaattttt gaaataaatg ggtggcttht aatggtgtct atgttaagtg aggctthtag 1800  
aatgctcttc ctgctthgtt tattatatgt gtatgaaaga tatgtatgta tttacatgtg 1860  
ttttagcgt cccagtcaa aacctgtgcg ctatacctaa atggattgat aatcttcatt 1920  
cactaattct aaaatagact tcttcccaa agaacggtgt aacgatgagg ctctatccag 1980  
ctgcttatct aaatcaactt taacgatgga tgatcttatg acacggggat ctttctthta 2040  
agttcttaga atttcagact gtaccgcagc tgatgaatca aacagcatta aaaagtgata 2100  
tgctcgaaaa tgtttttcct ggtctttctt cattattthta ggaagatacc ttatgcccac 2160  
gggtacaatg tccctcacca cacctctgtt ttgaataatc agtttccga ttgtggaaga 2220

caattctttt gcttccaact ttggcgcat tggagttggt atgcgaacaa gtccgatcag 2280  
ctcataaagc atcttagtga aaagggtggt tttgcgttat tctttcctct gttgaagctt 2340

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 401 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met	Ala	Asn	Pro	Phe	Ser	Arg	Trp	Phe	Leu	Ser	Glu	Arg	Pro	Pro
1				5					10					15
Asn	Cys	His	Val	Ala	Asp	Leu	Glu	Thr	Ser	Leu	Asp	Pro	His	Gln
				20					25					30
Thr	Leu	Leu	Lys	Val	Gln	Lys	Tyr	Lys	Pro	Ala	Leu	Ser	Asp	Trp
				35					40					45
Val	His	Tyr	Ile	Phe	Leu	Gly	Ser	Ile	Met	Leu	Phe	Val	Phe	Ile
				50					55					60
Thr	Asn	Pro	Ala	Pro	Trp	Ile	Phe	Lys	Ile	Leu	Phe	Tyr	Cys	Phe
				65					70					75
Leu	Gly	Thr	Leu	Phe	Ile	Ile	Pro	Ala	Thr	Ser	Gln	Phe	Phe	Phe
				80					85					90
Asn	Ala	Leu	Pro	Ile	Leu	Thr	Trp	Val	Ala	Leu	Tyr	Phe	Thr	Ser
				95					100					105
Ser	Tyr	Phe	Pro	Asp	Asp	Arg	Arg	Pro	Pro	Ile	Thr	Val	Lys	Val
				110					115					120
Leu	Pro	Ala	Val	Glu	Thr	Ile	Leu	Tyr	Gly	Asp	Asn	Leu	Ser	Asp
				125					130					135
Ile	Leu	Ala	Thr	Ser	Thr	Asn	Ser	Phe	Leu	Asp	Ile	Leu	Ala	Trp
				140					145					150

Leu Pro Tyr Gly	Leu Phe His Phe Gly	Ala Pro Phe Val Val	Ala
155	160	165	
Ala Ile Leu Phe Val	Phe Gly Pro Pro Thr Val	Leu Gln Gly Tyr	
170	175	180	
Ala Phe Ala Phe Gly	Tyr Met Asn Leu Phe Gly	Val Ile Met Gln	
185	190	195	
Asn Val Phe Pro Ala	Ala Pro Pro Trp Tyr Lys	Ile Leu Tyr Gly	
200	205	210	
Leu Gln Ser Ala Asn	Tyr Asp Met His Gly	Ser Pro Gly Gly	Leu
215	220	225	
Ala Arg Ile Asp Lys	Leu Leu Gly Ile Asn	Met Tyr Thr Thr	Ala
230	235	240	
Phe Ser Asn Ser Ser	Val Ile Phe Gly Ala	Phe Pro Ser Leu	His
245	250	255	
Ser Gly Cys Ala Thr	Met Glu Ala Leu Phe	Phe Cys Tyr Cys	Phe
260	265	270	
Pro Lys Leu Lys Pro	Leu Phe Ile Ala Tyr	Val Cys Trp Leu	Trp
275	280	285	
Trp Ser Thr Met Tyr	Leu Thr His His Tyr	Phe Val Asp Leu	Met
290	295	300	
Ala Gly Ser Val Leu	Ser Tyr Val Ile Phe	Gln Tyr Thr Lys	Tyr
305	310	315	
Thr His Leu Pro Ile	Val Asp Thr Ser Leu	Phe Cys Arg Trp	Ser
320	325	330	
Tyr Thr Ser Ile Glu	Lys Tyr Asp Ile Ser	Lys Ser Asp Pro	Leu
335	340	345	
Ala Ala Asp Ser Asn	Asp Ile Glu Ser Val	Pro Leu Ser Asn	Leu
350	355	360	
Glu Leu Asp Phe Asp	Leu Asn Met Thr Asp	Glu Pro Ser Val	Ser
365	370	375	
Pro Ser Leu Phe Asp	Gly Ser Thr Ser Val	Ser Arg Ser Ser	Ala
380	385	390	



Thr Ser Ile Thr Ser Leu Gly Val Lys Arg Ala

395

400

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5340 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
agcgcttcta ttttctccc caccgcgagg cggaaatggc acattttttt tcttttgctt 60
ctgtgctttt gctgtaattt ttggcatgtg ctattgtatg aagataacgc gtggttccgt 120
ggaaatagcc ggaaattttg ccgggaatat gacggacatg atttaacacc cgtggaaatg 180
aaaaaagcca aggtaagaaa gtggcaatat ttttctaca aatagatctg ctgtccctta 240
gatgattacc atacatatat atatttatta cacacatctg tcagaggtag ctagcgaagg 300
tgtcactgaa atattttttt ttccagttag tataaatacg gaggtagaac agctctccgc 360
gtgtatatct ttttttgctg tatakaagaa caggaagaac gcatttccat acctttttct 420
ccttacaggt gccctctgag tagtgtcacg aacgaggaaa aagattaata ttactgtttt 480
tatattcaaa aagagtaaag ccgttgctat atacgaatat gacgattacc gtgggggatg 540
cagtttcgga gacggagctg gaaaacaaaa gtcaaaacgt ggtactatct cccaaggcat 600
ctgcttcttc agacataagc acagatgttg ataaagacac atcgtcttct tgggatgaca 660
aatctttgct gcctacaggt gaatatattg tggacagaaa taagcccaa acctacttga 720
atagcgatga tatcgaaaaa gtgacagaat ctgatatttt ccctcagaaa cgtctgtttt 780
cattcttgca ctctaagaaa attccagaag taccacaaac cgatgacgag aggaagatat 840
atcctctgtt ccatacaaat attatctcta acatgttttt ttggtgggtt ctacccatcc 900
tgcgagtggg ttataagaga acgatacagc cgaacgatct cttcaaaatg gatccgagga 960
tgtctataga gaccctttat gacgactttg aaaaaaacat gatttactat tttgagaaga 1020
cgaggaaaaa ataccgtaaa agacatccag aagcgacaga agaagagggt atggaaaaatg 1080
ccaaactacc taaacataca gttctgagag ctttattatt cacttttaag aaacagtact 1140
tcatgtcgat agtgtttgca attctcgcta attgtacatc cggttttaac cccatgatta 1200
```

ccaagaggct aattgagttt gtcgaagaaa aggctatttt tcatagcatg catgttaaca 1260  
 aaggtattgg ttacgctatt ggtgcatggt tgatgatggt cgtaaagggt ttgacgttca 1320  
 atcatttctt tcatacatcc caactgactg gtgtgcaagc taagtctatt cttactaaag 1380  
 ctgccatgaa gaaaatgttt aatgcatcta attatgagag acattgtttt cctaacggta 1440  
 aagtgacttc ttttgtaaca acagatctcg ctagaattga atttgcctta tcttttcagc 1500  
 cgtttttggc tgggttcctt gcaattttgg ctatttgcag tgttttattg atcgtaaacc 1560  
 ttggacccat tgccttagtt gggattggta tttttttcgg tgggtttttc atatccttat 1620  
 ttgcatttaa gtttaattctg ggcttttagaa ttgctgcaaa catcttcact gatgctagag 1680  
 ttaccatgat gagagaagtg ctgaataata taaaaatgat taaatattat acgtgggagg 1740  
 atgcgtatga aaaaaatatt caagatatta ggaccaaaga gatttctaaa gttagaaaaa 1800  
 tgcaactatc aagaaatttc ttgattgcta tggccatgtc tttgcctagt attgcttcat 1860  
 tggtcacttt ccttgcaatg taaaaagtta ataaaggagg caggcaacct ggtaatatatt 1920  
 ttgcctcttt atctttattt caggctctga gtttgcaaat gtttttctta cctattgcta 1980  
 ttggtactgg aattgacatg atcattggat tgggccgttt gcaaagctta ttggaggctc 2040  
 cagaagatga tccaaatcag atgattgaaa tgaagccctc tcctggcttt gatccaaaat 2100  
 tggctctaaa aatgacacat tgctcatttg agtgggaaga ttatgaatta aacgacgcta 2160  
 ttgaagaagc aaaaggagaa gctaaagatg aaggtaaaaa gaacaaaaaa aagcgtaagg 2220  
 atacatgggg taagccatct gcaagtacta ataaggcgaa aagattggac aatatgttga 2280  
 aagacagaga cggcccgaa gatttagaaa aaacttcggt taggggtttc aaggacttga 2340  
 acttcgatat taaaaagggc gaatttatta tgattacggg acctattggg actggtaaat 2400  
 cttcattatt gaatgcatg gcaggatcaa tgagaaaaat tgatggtaag gttgaagtca 2460  
 acggggactt attaattgtg gggttatccat ggattcaaaa tgcattctga agagataaca 2520  
 tcatattcgg ttcaccattc aataaagaaa agtatgatga agtagttcgt gtttgctctt 2580  
 tgaaagctga tctggatatt ttaccggcag gcgatatgac cgaaattggg gaacgtggta 2640  
 ttactttatc tgggtgtcaa aaggcacgta tcaatttagc caggctctgt tataagaaga 2700  
 aggatattta tgtattcgac gatgtcctaa gtgctgtcga ttctcgtgtt ggtaaacaca 2760  
 tcatggatga atgtctaacc ggaatgcttg ctaataaaac cagaatttta gcaacgcatac 2820  
 agttgtcact gattgagaga gcttctagag tcatcggttt aggtactgat ggccaagtcg 2880  
 atattggtac tgttgatgag ctaaaagctc gtaatcaaac tttgataaat cttttacaat 2940  
 tctcttctca aaattcggag aaagaggatg aagaacagga agcggttggt tccggtgaat 3000  
 tgggacaact aaaatatgaa ccagaggtaa aggaattgac tgaactgaag aaaaaggcta 3060  
 cagaaatgtc acaaactgca aatagtggta aaattgtagc ggatgggtcat actagtagta 3120

aagaagaaag agcagtcaat agtatcagtc tgaaaatata ccgtgaatac attaaagctg 3180  
cagtaggtaa gtgggggtttt atcgccactac cggtgtatgc aatttttagtc gttggaacca 3240  
cattctgctc acttttttct tccgtttggt tatcttactg gactgagaat aaattcaaaa 3300  
acagaccacc cagtttttat atgggtcttt actccttctt tgtgtttgct gctttcatat 3360  
tcatgaatgg ccagttcacc atactttgcg caatgggtat tatggcatcg aaatgggttaa 3420  
atgtgagggc tgtgaaaaga attttacaca ctccaatgtc atacatagat accacacctt 3480  
tgggacgtat tctgaacaga ttcacaaaag atacagatag cttagataat gagttaaccg 3540  
aaagtttacg gttgatgaca tctcaatttg ctaatattgt aggtgtttgc gtcattgtga 3600  
ttgtttactt gccgtgggtt gctatcgcaa ttccgtttct tttggtcatc tttgttctga 3660  
ttgctgatca ttatcagagt tctggtagag aaattaaaag acttgaagct gtgcaacggg 3720  
cttttgttta caataattta aatgaagttt tgggtgggat ggatacaatc aaagcatacc 3780  
gaagtcagga acgatttttg gcgaaatcag attttttgat caacaagatg aatgaggcgg 3840  
gataccttgt agttgtcctg caaagatggg taggtatttt ccttgatatg gttgctatcg 3900  
catttgcact aattattacg ttattgtgtg ttacgagagc ctttcctatt tccgcggctt 3960  
cagttggtgt tttgttgact tatgtattac aattgcctgg tctattaaat accattttta 4020  
gggcaatgac tcaaacagag aatgacatga atagtgccga aagattggtg acatatgcaa 4080  
ctgaactacc actagaggca tcctatagaa agcccgaaat gacacctcca gagtcattggc 4140  
cctcaatggg cgaaataatt tttgaaaatg ttgattttgc ctatagacct gggtttaccta 4200  
tagtttttaa aaatcttaac ttgaatatca agagtgggga aaaaattggt atctgtggtc 4260  
gtacaggtgc tggtaatgcc actattatga gtgcccttta caggttgaat gaattgaccg 4320  
caggtaaaat ttaattgac aatgttgata taagtcagct gggacttttc gatttaagaa 4380  
gaaaattagc catcattcca caagatccag tattatttag gggtacgatt cgcaagaact 4440  
tagatccatt taatgagcgt acagatgacg aattatggga tgcattggtg agagggtgtg 4500  
ctatcgccaa ggatgacttg ccggaagtga aattgcaaaa acctgatgaa aatggtactc 4560  
atggtaaaat gcataagttc catttagatc aagcagtga agaagagggc tccaatttct 4620  
ccttaggtga gagacaacta ttagcattaa caagggcatt ggtccgcaa tcaaaaatat 4680  
tgattttgga tgaggctaca tcctcagtgg actacgaaac ggatggcaaa atccaaacac 4740  
gtattgttga ggaatttga gattgtacaa ttttgtgtat tgctcacaga ctgaagacca 4800  
ttgtaaatta tgatcgtatt cttgttttag agaaggggtga agtcgcagaa ttcgatacac 4860  
catggacgtt gtttagtcaa gaagatagta ttttcagaag catgtgttct agatctggta 4920  
ttgtggaaaa tgatttcgag aacagaagtt aatttatatt atttgttgca tgatttttct 4980  
ctttatttta tttatatgtt gccgatggta caaattagta ctagaaaaga aaaccacta 5040

ctatgacttg cagaaaaagt tatgtgtgcc atagatagat ataattgcat acccaccatcg 5100  
tataactcaaa attccgaaaa gaacatttca ttttttatga ggcaaactga acaacgcttc 5160  
ggtccttttt tcattctaga aatatatatt tatacatcat tttcagaaga tattcaaaga 5220  
acttattggg atgtctattt actgaataaa gtatacacia aaaacgaatt taaaatggaa 5280  
ggcataaata gaaaacttag aagtgaaaat cctaaaaccg aaggatattt caaatacgta 5340

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1477 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met	Thr	Ile	Thr	Val	Gly	Asp	Ala	Val	Ser	Glu	Thr	Glu	Leu	Glu
				5					10					15
Asn	Lys	Ser	Gln	Asn	Val	Val	Leu	Ser	Pro	Lys	Ala	Ser	Ala	Ser
				20					25					30
Ser	Asp	Ile	Ser	Thr	Asp	Val	Asp	Lys	Asp	Thr	Ser	Ser	Ser	Trp
				35					40					45
Asp	Asp	Lys	Ser	Leu	Leu	Pro	Thr	Gly	Glu	Tyr	Ile	Val	Asp	Arg
				50					55					60
Asn	Lys	Pro	Gln	Thr	Tyr	Leu	Asn	Ser	Asp	Asp	Ile	Glu	Lys	Val
				65					70					75
Thr	Glu	Ser	Asp	Ile	Phe	Pro	Gln	Lys	Arg	Leu	Phe	Ser	Phe	Leu
				80					85					90
His	Ser	Lys	Lys	Ile	Pro	Glu	Val	Pro	Gln	Thr	Asp	Asp	Glu	Arg
				95					100					105
Lys	Ile	Tyr	Pro	Leu	Phe	His	Thr	Asn	Ile	Ile	Ser	Asn	Met	Phe
				110					115					120
Phe	Trp	Trp	Val	Leu	Pro	Ile	Leu	Arg	Val	Gly	Tyr	Lys	Arg	Thr

125	130	135
Ile Gln Pro Asn Asp Leu Phe Lys Met	Asp Pro Arg Met Ser Ile	
140	145	150
Glu Thr Leu Tyr Asp Asp Phe Glu Lys	Asn Met Ile Tyr Tyr Phe	
155	160	165
Glu Lys Thr Arg Lys Lys Tyr Arg Lys	Arg His Pro Glu Ala Thr	
170	175	180
Glu Glu Glu Val Met Glu Asn Ala Lys	Leu Pro Lys His Thr Val	
185	190	195
Leu Arg Ala Leu Leu Phe Thr Phe Lys	Lys Gln Tyr Phe Met Ser	
200	205	210
Ile Val Phe Ala Ile Leu Ala Asn Cys	Thr Ser Gly Phe Asn Pro	
215	220	225
Met Ile Thr Lys Arg Leu Ile Glu Phe	Val Glu Glu Lys Ala Ile	
230	235	240
Phe His Ser Met His Val Asn Lys Gly	Ile Gly Tyr Ala Ile Gly	
245	250	255
Ala Cys Leu Met Met Phe Val Asn Gly	Leu Thr Phe Asn His Phe	
260	265	270
Phe His Thr Ser Gln Leu Thr Gly Val	Gln Ala Lys Ser Ile Leu	
275	280	285
Thr Lys Ala Ala Met Lys Lys Met Phe	Asn Ala Ser Asn Tyr Ala	
290	295	300
Arg His Cys Phe Pro Asn Gly Lys Val	Thr Ser Phe Val Thr Thr	
305	310	315
Asp Leu Ala Arg Ile Glu Phe Ala Leu	Ser Phe Gln Pro Phe Leu	
320	325	330
Ala Gly Phe Pro Ala Ile Leu Ala Ile	Cys Ile Val Leu Leu Ile	
335	340	345
Val Asn Leu Gly Pro Ile Ala Leu Val	Gly Ile Gly Ile Phe Phe	
350	355	360
Gly Gly Phe Phe Ile Ser Leu Phe Ala	Phe Lys Leu Ile Leu Gly	

365	370	375
Phe Arg Ile Ala	Ala Asn Ile Phe Thr Asp Ala Arg Val Thr Met	
380	385	390
Met Arg Glu Val	Leu Asn Asn Ile Lys Met Ile Lys Tyr Tyr Thr	
395	400	405
Trp Glu Asp Ala	Tyr Glu Lys Asn Ile Gln Asp Ile Arg Thr Lys	
410	415	420
Glu Ile Ser Lys	Val Arg Lys Met Gln Leu Ser Arg Asn Phe Leu	
425	430	435
Ile Ala Met Ala	Met Ser Leu Pro Ser Ile Ala Ser Leu Val Thr	
440	445	450
Phe Leu Ala Met	Tyr Lys Val Asn Lys Gly Gly Arg Gln Pro Gly	
455	460	465
Asn Ile Phe Ala	Ser Leu Ser Leu Phe Gln Val Leu Ser Leu Gln	
470	475	480
Met Phe Phe Leu	Pro Ile Ala Ile Gly Thr Gly Ile Asp Met Ile	
485	490	495
Ile Gly Leu Gly	Arg Leu Gln Ser Leu Leu Glu Ala Pro Glu Asp	
500	505	510
Asp Pro Asn Gln	Met Ile Glu Met Lys Pro Ser Pro Gly Phe Asp	
515	520	525
Pro Lys Leu Ala	Leu Lys Met Thr His Cys Ser Phe Glu Trp Glu	
530	535	540
Asp Tyr Glu Leu	Asn Asp Ala Ile Glu Glu Ala Lys Gly Glu Ala	
545	550	555
Lys Asp Glu Gly	Lys Lys Asn Lys Lys Lys Arg Lys Asp Thr Trp	
560	565	570
Gly Lys Pro Ser	Ala Ser Thr Asn Lys Ala Lys Arg Leu Asp Asn	
575	580	585
Met Leu Lys Asp	Arg Asp Gly Pro Glu Asp Leu Glu Lys Thr Ser	
590	595	600
Phe Arg Gly Phe	Lys Asp Leu Asn Phe Asp Ile Lys Lys Gly Glu	

605	610	615
Phe Ile Met Ile Thr Gly Pro Ile Gly Thr Gly Lys Ser Ser Leu		
620	625	630
Leu Asn Ala Met Ala Gly Ser Met Arg Lys Ile Asp Gly Lys Val		
635	640	645
Glu Val Asn Gly Asp Leu Leu Met Cys Gly Tyr Pro Trp Ile Gln		
650	655	660
Asn Ala Ser Val Arg Asp Asn Ile Ile Phe Gly Ser Pro Phe Asn		
665	670	675
Lys Glu Lys Tyr Asp Glu Val Val Arg Val Cys Ser Leu Lys Ala		
680	685	690
Asp Leu Asp Ile Leu Pro Ala Gly Asp Met Thr Glu Ile Gly Glu		
695	700	705
Arg Gly Ile Thr Leu Ser Gly Gly Gln Lys Ala Arg Ile Asn Leu		
710	715	720
Ala Arg Ser Val Tyr Lys Lys Lys Asp Ile Tyr Val Phe Asp Asp		
725	730	735
Val Leu Ser Ala Val Asp Ser Arg Val Gly Lys His Ile Met Asp		
740	745	750
Glu Cys Leu Thr Gly Met Leu Ala Asn Lys Thr Arg Ile Leu Ala		
755	760	765
Thr His Gln Leu Ser Leu Ile Glu Arg Ala Ser Arg Val Ile Val		
770	775	780
Leu Gly Thr Asp Gly Gln Val Asp Ile Gly Thr Val Asp Glu Leu		
785	790	795
Lys Ala Arg Asn Gln Thr Leu Ile Asn Leu Leu Gln Phe Ser Ser		
800	805	810
Gln Asn Ser Glu Lys Glu Asp Glu Glu Gln Glu Ala Val Val Ser		
815	820	825
Gly Glu Leu Gly Gln Leu Lys Tyr Glu Pro Glu Val Lys Glu Leu		
830	835	840
Thr Glu Leu Lys Lys Lys Ala Thr Glu Met Ser Gln Thr Ala Asn		

845	850	855
Ser Gly Lys Ile Val Ala Asp Gly His Thr Ser Ser Lys Glu Glu		
860	865	870
Arg Ala Val Asn Ser Ile Ser Leu Lys Ile Tyr Arg Glu Tyr Ile		
875	880	885
Lys Ala Ala Val Gly Lys Trp Gly Phe Ile Ala Leu Pro Leu Tyr		
890	895	900
Ala Ile Leu Val Val Gly Thr Thr Phe Cys Ser Leu Phe Ser Ser		
905	910	915
Val Trp Leu Ser Tyr Trp Thr Glu Asn Lys Phe Lys Asn Arg Pro		
920	925	930
Pro Ser Phe Tyr Met Gly Leu Tyr Ser Phe Phe Val Phe Ala Ala		
935	940	945
Phe Ile Phe Met Asn Gly Gln Phe Thr Ile Leu Cys Ala Met Gly		
950	955	960
Ile Met Ala Ser Lys Trp Leu Asn Leu Arg Ala Val Lys Arg Ile		
965	970	975
Leu His Thr Pro Met Ser Tyr Ile Asp Thr Thr Pro Leu Gly Arg		
980	985	990
Ile Leu Asn Arg Phe Thr Lys Asp Thr Asp Ser Leu Asp Asn Glu		
995	1000	1005
Leu Thr Glu Ser Leu Arg Leu Met Thr Ser Gln Phe Ala Asn Ile		
1010	1015	1020
Val Gly Val Cys Val Met Cys Ile Val Tyr Leu Pro Trp Phe Ala		
1025	1030	1035
Ile Ala Ile Pro Phe Leu Leu Val Ile Phe Val Leu Ile Ala Asp		
1040	1045	1050
His Tyr Gln Ser Ser Gly Arg Glu Ile Lys Arg Leu Glu Ala Val		
1055	1060	1065
Gln Arg Ser Phe Val Tyr Asn Asn Leu Asn Glu Val Leu Gly Gly		
1070	1075	1080
Met Asp Thr Ile Lys Ala Tyr Arg Ser Gln Glu Arg Phe Leu Ala		



1085	1090	1095
Lys Ser Asp Phe Leu Ile Asn Lys Met Asn Glu Ala Gly Tyr Leu		
1100	1105	1110
Val Val Val Leu Gln Arg Trp Val Gly Ile Phe Leu Asp Met Val		
1115	1120	1125
Ala Ile Ala Phe Ala Leu Ile Ile Thr Leu Leu Cys Val Thr Arg		
1130	1135	1140
Ala Phe Pro Ile Ser Ala Ala Ser Val Gly Val Leu Leu Thr Tyr		
1145	1150	1155
Val Leu Gln Leu Pro Gly Leu Leu Asn Thr Ile Leu Arg Ala Met		
1160	1165	1170
Thr Gln Thr Glu Asn Asp Met Asn Ser Ala Glu Arg Leu Val Thr		
1175	1180	1185
Tyr Ala Thr Glu Leu Pro Leu Glu Ala Ser Tyr Arg Lys Pro Glu		
1190	1195	1200
Met Thr Pro Pro Glu Ser Trp Pro Ser Met Gly Glu Ile Ile Phe		
1205	1210	1215
Glu Asn Val Asp Phe Ala Tyr Arg Pro Gly Leu Pro Ile Val Leu		
1220	1225	1230
Lys Asn Leu Asn Leu Asn Ile Lys Ser Gly Glu Lys Ile Gly Ile		
1235	1240	1245
Cys Gly Arg Thr Gly Ala Gly Lys Ser Thr Ile Met Ser Ala Leu		
1250	1255	1260
Tyr Arg Leu Asn Glu Leu Thr Ala Gly Lys Ile Leu Ile Asp Asn		
1265	1270	1275
Val Asp Ile Ser Gln Leu Gly Leu Phe Asp Leu Arg Arg Lys Leu		
1280	1285	1290
Ala Ile Ile Pro Gln Asp Pro Val Leu Phe Arg Gly Thr Ile Arg		
1295	1300	1305
Lys Asn Leu Asp Pro Phe Asn Glu Arg Thr Asp Asp Glu Leu Trp		
1310	1315	1320
Asp Ala Leu Val Arg Gly Gly Ala Ile Ala Lys Asp Asp Leu Pro		

1325	1330	1335
Glu Val Lys Leu Gln Lys Pro Asp Glu Asn Gly Thr His Gly Lys		
1340	1345	1350
Met His Lys Phe His Leu Asp Gln Ala Val Glu Glu Glu Gly Ser		
1355	1360	1365
Asn Phe Ser Leu Gly Glu Arg Gln Leu Leu Ala Leu Thr Arg Ala		
1370	1375	1380
Leu Val Arg Gln Ser Lys Ile Leu Ile Leu Asp Glu Ala Thr Ser		
1385	1390	1395
Ser Val Asp Tyr Glu Thr Asp Gly Lys Ile Gln Thr Arg Ile Val		
1400	1405	1410
Glu Glu Phe Gly Asp Cys Thr Ile Leu Cys Ile Ala His Arg Leu		
1415	1420	1425
Lys Thr Ile Val Asn Tyr Asp Arg Ile Leu Val Leu Glu Lys Gly		
1430	1435	1440
Glu Val Ala Glu Phe Asp Thr Pro Trp Thr Leu Phe Ser Gln Glu		
1445	1450	1455
Asp Ser Ile Phe Arg Ser Met Cys Ser Arg Ser Gly Ile Val Glu		
1460	1465	1470
Asn Asp Phe Glu Asn Arg Ser		
1475		

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

tttggttaya tgaayytntt yggngt 26

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

tctacaaart artggtgngt narrtacat 29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2274 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
ttatatatat tattgatttg ttcctgttgt tatttagttt agaatcagac gactacacca 60
gaaccacaat tcaaccaaca cttatataga acctggcttg gaaaaaagta acatttatca 120
ttcctatact tttttagcaa acataatccg tgttttacat atattattca cccaatatca 180
taacaaaaac aaactgaata atggcgtcct ctattttgcg ttccaaaata atacaaaaac 240
cgtaccaatt attccactac tattttcttc tggagaaggc tcctgggttct acagttagtg 300
atgtgaattt tgatacaaac atacaaacga gtttacgtaa attaaagcat catcattgga 360
cgggtgggaga aatattccat tatggggttt tggtttccat actttttttc gtgtttgtgg 420
ttttcccagc ttcatttttt ataaaattac caataatctt agcatttgct acttgttttt 480
taataccctt aacatcacaa tttttcttc ctgccttgcc cgttttcact tggttggcat 540
tatattttac gtgtgctaaa atacctcaag aatggaaacc agctatcaca gttaaagttt 600
taccagctat ggaaacaatt ttgtacggcg ataatttatc aaatgttttg gcaaccatca 660
```

```

ctaccggagt gttagatata ttggcatggg taccatatgg gattattcat ttcagtttcc 720
catttgact tgctgtatt atatTTTTat ttgggccacc gacggcatta agatcatttg 780
gatttgcctt tggttatatg aacttgcttg gagtcttgat tcaaatggca ttcccagctg 840
ctcctccatg gtacaaaaac ttgcacggat tagaaccagc taattattca atgcacgggt 900
ctcctggtgg acttgaaggg atagataaat tgttaggtgt tgatatgtat accacagggg 960
tttccaattc atcaatcatt tttggggcat tcccatcgtt acattcagga tgttgtatca 1020
tggaagtgtt atttttgtgt tggttgtttc cacgattcaa gtttgtgtgg gttacatacg 1080
catcttggct ttggtggagc acgatgtatt tgacctatca ctactttgtc gatttgattg 1140
gtggagccat gctatctttg actgtttttg agttcaccaa atataaatat ttgccaaaaa 1200
acaaagaagg ctttttctgt cgttgggtcat aactgaaat tgaaaaaatc gatatccaag 1260
agattgacct tttatcatac aattatatcc ctgtcaacag caatgataat gaaagcagat 1320
tgtatacgag agtgtaccaa gagtctcagg ttagtcccc acagagagct gaaacacctg 1380
aagcatttga gatgtcaaat ttttctaggt ctagacaaag ctcaaagact caggttccat 1440
tgagtaatct tactaacaat gatcaagtgt ctggaattaa cgaagaggat gaagaagaag 1500
aaggcgatga aatttcatcg agtactcctt cgggtgtttga agacgaacca cagggttagca 1560
catatgctgc atcctcagct acatcagtag atgatttgga ttccaaaaga aattagtaaa 1620
ataacagttt ctattaattt ctttatttcc tcctaattaa tgattttatg ctcaatacct 1680
acactatctg tttttaattt cctacttttt ttttattatt gttgagttca tttgctgttc 1740
attgaatatt tacaattttg cattaattac catcaatata gaatgggcac agttttttta 1800
agtttttttg tttttgtgtt tgtctttctt tttttacatt aatgtgtttg gattgtttta 1860
ggttccttta tcccttagcc ccctcagaat actatTTTTat ctaattaatt tgtttttatt 1920
ttctgatatt taccaattgc ttttctttt ggatatttat aatagcatcc cctaataatt 1980
aatatacaac tgtttcatat atatacgtgt atgtcctgta gtggtggaaa ctggagtcaa 2040
catttgattt aatgtgtaca agaaagcagt gttaatgcta ctattataat ttttgagggtg 2100
caaatcaaga ggttggcagc tttcttatgg ctatgaccgt gaatgaaggc ttgtaaacca 2160
cgtaataaac aaaagccaac aagttttttt agagccttta acaacatacg caatgagagt 2220
gattgcaata ctacaagata tagcccaaaa aattgaatgc atttcaacaa caac 2274

```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 471 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met	Ala	Ser	Ser	Ile	Leu	Arg	Ser	Lys	Ile	Ile	Gln	Lys	Pro	Tyr
				5					10					15
Gln	Leu	Phe	His	Tyr	Tyr	Phe	Leu	Ser	Glu	Lys	Ala	Pro	Gly	Ser
				20					25					30
Thr	Val	Ser	Asp	Leu	Asn	Phe	Asp	Thr	Asn	Ile	Gln	Thr	Ser	Leu
				35					40					45
Arg	Lys	Leu	Lys	His	His	His	Trp	Thr	Val	Gly	Glu	Ile	Phe	His
				50					55					60
Tyr	Gly	Phe	Leu	Val	Ser	Ile	Leu	Phe	Phe	Val	Phe	Val	Val	Phe
				65					70					75
Pro	Ala	Ser	Phe	Phe	Ile	Lys	Leu	Pro	Ile	Ile	Leu	Ala	Phe	Ala
				80					85					90
Thr	Cys	Phe	Leu	Ile	Pro	Leu	Thr	Ser	Gln	Phe	Phe	Leu	Pro	Ala
				95					100					105
Leu	Pro	Val	Phe	Thr	Trp	Leu	Ala	Leu	Tyr	Phe	Thr	Cys	Ala	Lys
				110					115					120
Ile	Pro	Gln	Glu	Trp	Lys	Pro	Ala	Ile	Thr	Val	Lys	Val	Leu	Pro
				125					130					135
Ala	Met	Glu	Thr	Ile	Leu	Tyr	Gly	Asp	Asn	Leu	Ser	Asn	Val	Leu
				140					145					150
Ala	Thr	Ile	Thr	Thr	Gly	Val	Leu	Asp	Ile	Leu	Ala	Trp	Leu	Pro
				155					160					165
Tyr	Gly	Ile	Ile	His	Phe	Ser	Phe	Pro	Phe	Val	Leu	Ala	Ala	Ile
				170					175					180
Ile	Phe	Leu	Phe	Gly	Pro	Pro	Thr	Ala	Leu	Arg	Ser	Phe	Gly	Phe
				185					190					195
Ala	Phe	Gly	Tyr	Met	Asn	Leu	Leu	Gly	Val	Leu	Ile	Gln	Met	Ala

200	205	210
Phe Pro Ala Ala Pro Pro Trp Tyr Lys	Asn Leu His Gly Leu Glu	
215	220	225
Pro Ala Asn Tyr Ser Met His Gly Ser	Pro Gly Gly Leu Gly Arg	
230	235	240
Ile Asp Lys Leu Leu Gly Val Asp Met	Tyr Thr Thr Gly Phe Ser	
245	250	255
Asn Ser Ser Ile Ile Phe Gly Ala Phe	Pro Ser Leu His Ser Gly	
260	265	270
Cys Cys Ile Met Glu Val Leu Phe Leu	Cys Trp Leu Phe Pro Arg	
275	280	285
Phe Lys Phe Val Trp Val Thr Tyr Ala	Ser Trp Leu Trp Trp Ser	
290	295	300
Thr Met Tyr Leu Thr His His Tyr Phe	Val Asp Leu Ile Gly Gly	
305	310	315
Ala Met Leu Ser Leu Thr Val Phe Glu	Phe Thr Lys Tyr Lys Tyr	
320	325	330
Leu Pro Lys Asn Lys Glu Gly Leu Phe	Cys Arg Trp Ser Tyr Thr	
335	340	345
Glu Ile Glu Lys Ile Asp Ile Gln Glu	Ile Asp Pro Leu Ser Tyr	
350	355	360
Asn Tyr Ile Pro Val Asn Ser Asn Asp	Asn Glu Ser Arg Leu Tyr	
365	370	375
Thr Arg Val Tyr Gln Glu Ser Gln Val	Ser Pro Pro Gln Arg Ala	
380	385	390
Glu Thr Pro Glu Ala Phe Glu Met Ser	Asn Phe Ser Arg Ser Arg	
395	400	405
Gln Ser Ser Lys Thr Gln Val Pro Leu	Ser Asn Leu Thr Asn Asn	
410	415	420
Asp Gln Val Ser Gly Ile Asn Glu Glu	Asp Glu Glu Glu Glu Gly	
425	430	435
Asp Glu Ile Ser Ser Ser Thr Pro Ser	Val Phe Glu Asp Glu Pro	

440	445	450
Gln Gly Ser Thr Tyr Ala Ala Ser Ser Ala Thr Ser Val Asp Asp		
455	460	465
Leu Asp Ser Lys Arg Asn		
470		

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 243 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```

tttgaaaaat ttgaatttta aaattaatcc aatggaaaaa attggtattt gtggaagaac   60
cggtgctggt aaatcatcaa ttatgacagc attatatcga ttatcagaat tagaactggg   120
gaaaattatt attgatgata ttgatatttc aactttgggt ttaaaagatc ttcgatcaaa   180
attatcaatt attcctcaag atccagtatt attccgaggt tcaattcgga aaaacttgga   240
tcc                                                                243

```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Leu Lys Asn Leu Asn Phe Lys Ile Asn Pro Met Glu Lys Ile Gly
5 10 15
55

Ile	Cys	Gly	Arg	Thr	Gly	Ala	Gly	Lys	Ser	Ser	Ile	Met	Thr	Ala
Leu	Tyr	Arg	Leu	Ser	Glu	Leu	Glu	Leu	Gly	Lys	Ile	Ile	Ile	Asp
Asp	Ile	Asp	Ile	Ser	Thr	Leu	Gly	Leu	Lys	Asp	Leu	Arg	Ser	Lys
Leu	Ser	Ile	Ile	Pro	Gln	Asp	Pro	Val	Leu	Phe	Arg	Gly	Ser	Ile
Arg	Lys	Asn	Leu	Asp										

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1601 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iv) ANTI-SENSE: Yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

aggaagatga cttgcatcaa agatggagga agtgggtactg gcaggacgat caatcaaadc	60
agcagcagga ctaggtaacg gctcaggtga tgatgaaccc acggaccatt catgatcggc	120
gtagcaagt tccatattgt taagaccact catgaaggct actgcattag ggttttgagt	180
aaaagaatcc cttccaagta agtatgggct gccggtacga gccaaggagt tgctgggttt	240
ttcggaaga ccatgaccgt ggataacaaa ctcgtattcc caacgaagga ttttaccagt	300
ttgcaactgt gggaggcgta gcttttgagc aaaaacgaag catataatag ctaaacacat	360
accgccgacc aaatctacaa agtagtggtg ggtaagggtac atagtacacc agcaaagcca	420
tagaacatat ccataaaagc agaagcggtg tcgaggaaac acatgcgaaa ggaaaagtgc	480
ttccagcatg gcccatccag cgtgaagaga tggaaaggca ccaaaaacaa ccggagagtt	540
agaaaaacca tcagtgtaaa tgctagtgcc gaagagagca tcaatacggg ccaatccacc	600
aggagagcca cgtactgcat acgtggcagg ttctaaacca tacatatattt cataccaagg	660



aggagaacag	gggaaagcca	tttggataag	aacaccaa	aaattcatat	aacccaaaagt	720
tcgagcccaa	actggaagag	ttccaggagg	tgcaaagatg	aaaagaataa	atgaaatgat	780
aaaaggagcc	gaataatgca	tgactccata	tggaacccag	gccaaaatat	caaggatgct	840
atgcgtggtt	ttcgagagaa	gactagaaag	attagagcca	taaagaatat	tttcaagtgt	900
gggtaaaaca	cgaacccata	tgggtggacg	ccagcgttct	ggaataaacc	tacaagagta	960
aaataaaatt	gcccagggtga	tgataacaat	ggcaggaaaa	aaaatttggc	gtgttaaagg	1020
aacggtcaac	gcaatggcca	aaagacaggc	aatgccaaat	ttccccaga	atccaggaga	1080
ttcaatgaca	atacaagcaa	aaatcaaatt	acctgctaga	aacacatat	gcaaatgtgt	1140
ccatgaccat	ttcgtattgc	gtagcaaacg	aaatgtaggc	atagggttta	agcttgtttc	1200
caacttgat	tgggatgctc	ggttacacgc	agcaaggcgc	ttttttaagg	tcgaaagagc	1260
agacattgct	tcaaagaatt	atcagagtaa	aaaagggag	cgtacgaaaa	aaatttcgta	1320
aggaattaac	cggaaaacta	aaggaaaaaa	aaggaatttt	tatgaaggaa	agaaagtagc	1380
tattaaatgc	aagtgtcaag	cacttaaaag	tagcgatgta	aaatatttaa	aaaaagatgg	1440
accgattaac	caatgttcag	ctcacagttg	ccagcaatca	gggctatttt	tttatttttt	1500
ttataaaaatt	gctaattata	tataatataa	ttagtttatt	aacttgcttt	tcctcaaaaa	1560
accaattcga	gaaaggaact	tttgagagg	caaaaaagct	t		1601

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1601 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: mRNA

(iv) ANTI-SENSE: Yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

aggaagauga	cuugcauc	aaagauggagga	agugguacug	gcaggacgau	caaucaaauc	60
agcagcagga	cuagguaacg	gcucagguga	ugaugaaccc	acggaccuu	caugaucggu	120
guuagcaagu	uccauauugu	uaagaccacu	caugaaggcu	acugcauuag	gguuuugagu	180
aaaagaaucc	cuuccaagua	aguaugggcu	gccgguacga	gccaaggagu	ugcugguuuu	240
uucggaaaga	ccaugaccgu	ggauaacaaa	cucguauucc	caacgaagga	uuuuaccagu	300

```

uugcaacugu gggaggcgua gcuuuugagc aaaaacgaag cauauauag cuaaacacau 360
accgccgacc aaaucuacaa aguaguggug gguaagguac auaguacacc agcaaagcca 420
uagaacauau ccuaaaaagc agaagcggua ucgaggaaac acaugcgaaa ggaaaagugc 480
uuccagcaug gcccauccag cgugaagaga uggaaaggca ccaaaaacaa ccggagaguu 540
agaaaaacca ucaguguaaa ugcuaugucc gaagagagca ucaauacggg ccaauccacc 600
aggagagcca cguacugcau acguggcagg uucuaaacca uacauauuuu cauaccaagg 660
aggagaacag gggaaagcca uuuggauaag aacaccaaau aaauucauau aaccaaagu 720
ucgagcccaa acuggaagag uuccaggagg ugcaaagaug aaaagaauaa augaaaugau 780
aaaaggagcc gaauaaugca ugacuccaua uggaaccag gccaaaauau caaggauucu 840
augcgugguu uucgagagaa gacuagaaag auuagagcca uaaagaauau uuucaagugu 900
ggguaaaaca cgaacccaua uggguggacg ccagcguucu ggaauaaacc uacaagagua 960
aaauaaaauu gcccagguga ugauaacaau ggcaggaaaa aaaauuuggc guguaaagg 1020
aacggucaac gcaauggcca aaagacaggc aaugccaaau uucccccaga auccaggaga 1080
uucaaugaca auacaagcaa aaaucaauu accugcuaga aacacauauu gcaaauugu 1140
ccaugaccau uucguauugc guagcaaacg aauguaggc auaggguuuu agcuuguuuc 1200
caacuuguau uggaugcuc gguuacacgc agcaaggcgc uuuuuuaagg ucgaaagagc 1260
agacauugcu ucaaagaauu aucagaguaa aaaagggaag cguacgaaaa aaauuucgua 1320
aggaauuaac cggaaaacua aaggaaaaaa aaggauuuu uaugaaggaa agaaaguagc 1380
uauuaaauugc aagugucaag cacuaaaaag uagcgaugua aaauauuuua aaaaagaugg 1440
accgauuaac cauguucag cucacaguug ccagcaauca gggcuauuuu uuuaauuuuu 1500
uuauaaaauu gcuaauuua uauaauuaa uuaguuuuu aacuugcuuu uccucaaaaa 1560
accaauucga gaaaggaacu uuugcagagg caaaaaagcu u 1601

```

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Cys Phe Thr Ser Ser Tyr Phe Pro Asp Asp Arg Arg  
5 10

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Cys Tyr Thr Ser Ile Glu Lys Tyr Asp Ile Ser Lys Ser Asp Pro  
5 10 15

Leu Ala Ala Asp

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1553 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```

ttttacatat attattcacc caatatcata acaaaaacaa actgaatgat ggcattcttct 60
at ttgtgcgtt ccaaaaataat acaaaaaccg taccaattat tccactacta ttttcttctg 120
gagaaggctc ctggttctac agttagtgat ttgaattttg atacaaacat acaaacgagt 180
ttacgtaaat taaagcatca tcattggacg gtgggagaaa tattccatta tgggtttttg 240
gtttccatac tttttttcgt gtttgtggtt ttcccagctt cattttttat aaaattacca 300
ataatcttag catttgctac ttgtttttta atacctttaa catcacaatt ttttcttcct 360
gccttgcccg ttttcacttg gttggcatta tatttttacgt gtgctaaaat acctcaagaa 420

```

```

tggaaccag ctatcacagt taaagtttta ccagctatgg aaacaatttt gtacggcgat 480
aatttatcaa atgttttggc aaccatcact accggagtgt tagatatatt ggcatgggta 540
ccatatggga ttattcattt cagtttccca tttgtacttg ctgctattat atttttattt 600
gggccaccga cggcattaag atcatttggg tttgcctttg gttatatgaa cttgcttgga 660
gtcttgattc aaatggcatt cccagctgct cctccatggg acaaaaaactt gcacggatta 720
gaaccagcta attattcaat gcacgggtct cctgggtggac ttggaaggat agataaattg 780
ttaggtgttg atatgtatac cacagggttt tccaattcat caatcatttt tggggcattc 840
ccatcgttac attcaggatg ttgtatcatg gaagtgttat ttttgtgttg gttgtttcca 900
cgattcaagt ttgtgtgggt tacatacgca tcttggcttt ggtggagcac gatgtatttg 960
acccatcact actttgtcga tttgattggg ggagccatgc tatctttgac tgtttttgaa 1020
ttcaccaaat ataaatattt gccaaaaaac aaagaaggcc ttttctgtcg ttggtcatac 1080
actgaaattg aaaaaatcga tatccaagag attgaccctt tatcatacaa ttatatccct 1140
gtcaacagca atgataatga aagcagattg tatacgagag tgtaccaaga gcctcagggt 1200
agtccccac agagagctga aacacctgaa gcatttgaga tgtcaaattt ttctagggtc 1260
agacaaagct caaagactca ggttccattg agtaatctta ctaacaatga tcaagtgcct 1320
ggaattaacg aagaggatga agaagaagaa ggcgatgaaa tttcgtcgag tactccttcg 1380
gtgtttgaag acgaaccaca gggtagcaca tatgctgcat cctcagctac atcagtagat 1440
gatttgatt ccaaagaaa ttagtaaaac agcagtttct attaatcttct ttatttcctc 1500
ctaattaatg attttatgtt caatacctac actatctgtt ttaatttcc tac 1553

```

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 472 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Met Met Ala Ser Ser Ile Leu Arg Ser Lys Ile Ile Gln Lys Pro

1 5 10 15

Tyr Gln Leu Phe His Tyr Tyr Phe Leu Leu Glu Lys Ala Pro Gly

60

	20	25	30
Ser Thr Val	Ser Asp Leu Asn Phe Asp Thr Asn Ile Gln Thr Ser		
	35	40	45
Leu Arg Lys Leu Lys His His His Trp Thr Val Gly Glu Ile Phe			
	50	55	60
His Tyr Gly Phe Leu Val Ser Ile Leu Phe Phe Val Phe Val Val			
	65	70	75
Phe Pro Ala Ser Phe Phe Ile Lys Leu Pro Ile Ile Leu Ala Phe			
	80	85	90
Ala Thr Cys Phe Leu Ile Pro Leu Thr Ser Gln Phe Phe Leu Pro			
	95	100	105
Ala Leu Pro Val Phe Thr Trp Leu Ala Leu Tyr Phe Thr Cys Ala			
	110	115	120
Lys Ile Pro Gln Glu Trp Lys Pro Ala Ile Thr Val Lys Val Leu			
	125	130	135
Pro Ala Met Glu Thr Ile Leu Tyr Gly Asp Asn Leu Ser Asn Val			
	140	145	150
Leu Ala Thr Ile Thr Thr Gly Val Leu Asp Ile Leu Ala Trp Leu			
	155	160	165
Pro Tyr Gly Ile Ile His Phe Ser Phe Pro Phe Val Leu Ala Ala			
	170	175	180
Ile Ile Phe Leu Phe Gly Pro Pro Thr Ala Leu Arg Ser Phe Gly			
	185	190	195
Phe Ala Phe Gly Tyr Met Asn Leu Leu Gly Val Leu Ile Gln Met			
	200	205	210
Ala Phe Pro Ala Ala Pro Pro Trp Tyr Lys Asn Leu His Gly Leu			
	215	220	225
Glu Pro Ala Asn Tyr Ser Met His Gly Ser Pro Gly Gly Leu Gly			
	230	235	240
Arg Ile Asp Lys Leu Leu Gly Val Asp Met Tyr Thr Thr Gly Phe			
	245	250	255
Ser Asn Ser Ser Ile Ile Phe Gly Ala Phe Pro Ser Leu His Ser			

260	265	270
Gly Cys Cys Ile Met Glu Val Leu Phe	Leu Cys Trp Leu Phe	Pro
275	280	285
Arg Phe Lys Phe Val Trp Val Thr Tyr	Ala Ser Trp Leu Trp	Trp
290	295	300
Ser Thr Met Tyr Leu Thr His His Tyr	Phe Val Asp Leu Ile	Gly
305	310	315
Gly Ala Met Leu Ser Leu Thr Val Phe	Glu Phe Thr Lys Tyr	Lys
320	325	330
Tyr Leu Pro Lys Asn Lys Glu Gly Leu	Phe Cys Arg Trp Ser	Tyr
335	340	345
Thr Glu Ile Glu Lys Ile Asp Ile Gln	Glu Ile Asp Pro Leu	Ser
350	355	360
Tyr Asn Tyr Ile Pro Val Asn Ser Asn	Asp Asn Glu Ser Arg	Leu
365	370	375
Tyr Thr Arg Val Tyr Gln Glu Pro Gln	Val Ser Pro Pro Gln	Arg
380	385	390
Ala Glu Thr Pro Glu Ala Phe Glu Met	Ser Asn Phe Ser Arg	Ser
395	400	405
Arg Gln Ser Ser Lys Thr Gln Val Pro	Leu Ser Asn Leu Thr	Asn
410	415	420
Asn Asp Gln Val Pro Gly Ile Asn Glu	Glu Asp Glu Glu Glu	Glu
425	430	435
Gly Asp Glu Ile Ser Ser Ser Thr Pro	Ser Val Phe Glu Asp	Glu
440	445	450
Pro Gln Gly Ser Thr Tyr Ala Ala Ser	Ser Ala Thr Ser Val	Asp
455	460	465
Asp Leu Asp Ser Lys Arg Asn		
470		

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

gactatttca ttatggggcc cc 22

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 bases
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

gttaactcga gaaagtgcc atcagtgttc 30

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 bases
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

gttaacggta cctcatcggt acaccgttc 29

(2) INFORMATION FOR SEQ ID NO:40:

- gctaaacgac aatcctgac 19

- cgttggccga ttcattaatg c 21

- Met Ala Asn Pro Phe Ser Arg Trp Phe Leu Ser Glu Arg Pro Pro

64



Asn Cys His Val Ala Asp Leu Glu Thr Ser Leu Asp Pro His Gln		
	20	25 30
Thr Leu Leu Lys Val Gln Lys Tyr Lys Pro Ala Leu Ser Asp Trp		
	35	40 45
Val His Tyr Ile Phe Leu Gly Ser Ile Met Leu Phe Val Phe Ile		
	50	55 60
Thr Asn Pro Ala Pro Trp Ile Phe Lys Ile Leu Phe Tyr Cys Phe		
	65	70 75
Leu Gly Thr Leu Phe Ile Ile Pro Ala Thr Ser Gln Phe Phe Phe		
	80	85 90
Asn Ala Leu Pro Ile Leu Thr Trp Val Ala Leu Tyr Phe Thr Ser		
	95	100 105
Ser Tyr Phe Pro Asp Asp Arg Arg Pro Pro Ile Thr Val Lys Val		
	110	115 120
Leu Pro Ala Val Glu Thr Ile Leu Tyr Gly Asp Asn Leu Ser Asp		
	125	130 135
Ile Leu Ala Thr Ser Thr Asn Ser Phe Leu Asp Ile Leu Ala Trp		
	140	145 150
Leu Pro Tyr Gly Leu Phe His Phe Gly Ala Pro Phe Val Val Ala		
	155	160 165
Ala Ile Leu Phe Val Phe Gly Pro Pro Thr Val Leu Gln Gly Tyr		
	170	175 180
Ala Phe Ala Phe Gly Tyr Met Asn Leu Phe Gly Val Ile Met Gln		
	185	190 195
Asn Val Phe Pro Ala Ala Pro Pro Trp Tyr Lys Ile Leu Tyr Gly		
	200	205 210
Leu Gln Ser Ala Asn Tyr Asp Met His Gly Ser Pro Gly Gly Leu		
	215	220 225
Ala Arg Ile Asp Lys Leu Leu Gly Ile Asn Met Tyr Thr Thr Cys		
	230	235 240
Phe Ser Asn Ser Ser Val Ile Phe Gly Ala Phe Pro Ser Leu His		
	245	250 255

Ser Gly Cys Ala Thr Met Glu Ala Leu Phe Phe Cys Tyr Cys Phe		
260	265	270
Pro Lys Leu Lys Pro Leu Phe Ile Ala Tyr Val Cys Trp Leu Trp		
275	280	285
Trp Ser Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Met		
290	295	300
Ala Gly Ser Val Leu Ser Tyr Val Ile Phe Gln Tyr Thr Lys Tyr		
305	310	315
Thr His Leu Pro Ile Val Asp Thr Ser Leu Phe Cys Arg Trp Ser		
320	325	330
Tyr Thr Ser Ile Glu Lys Tyr Asp Ile Ser Lys Ser Asp Pro Leu		
335	340	345
Ala Ala Asp Ser Asn Asp Ile Glu Ser Val Pro Leu Ser Asn Leu		
350	355	360
Glu Leu Asp Phe Asp Leu Asn Met Thr Asp Glu Pro Ser Val Ser		
365	370	375
Pro Ser Leu Phe Asp Gly Ser Thr Ser Val Ser Arg Ser Ser Ala		
380	385	390
Thr Ser Ile Thr Ser Leu Gly Val Lys Arg Ala		
395	400	

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 401 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Met Ala Asn Pro Phe Ser Arg Trp Phe Leu Ser Glu Arg Pro Pro			
1	5	10	15
		66	

Asn Cys His Val Ala Asp Leu Glu Thr Ser Leu Asp Pro His Gln		
	20	25 30
Thr Leu Leu Lys Val Gln Lys Tyr Lys Pro Ala Leu Ser Asp Trp		
	35	40 45
Val His Tyr Ile Phe Leu Gly Ser Ile Met Leu Phe Val Phe Ile		
	50	55 60
Thr Asn Pro Ala Pro Trp Ile Phe Lys Ile Leu Phe Tyr Cys Phe		
	65	70 75
Leu Gly Thr Leu Phe Ile Ile Pro Ala Thr Ser Gln Phe Phe Phe		
	80	85 90
Asn Ala Leu Pro Ile Leu Thr Trp Val Ala Leu Tyr Phe Thr Ser		
	95	100 105
Ser Tyr Phe Pro Asp Asp Arg Arg Pro Pro Ile Thr Val Lys Val		
	110	115 120
Leu Pro Ala Val Glu Thr Ile Leu Tyr Gly Asp Asn Leu Ser Asp		
	125	130 135
Ile Leu Ala Thr Ser Thr Asn Ser Phe Leu Asp Ile Leu Ala Trp		
	140	145 150
Leu Pro Tyr Gly Leu Phe His Tyr Gly Ala Pro Phe Val Val Ala		
	155	160 165
Ala Ile Leu Phe Val Phe Gly Pro Pro Thr Val Leu Gln Gly Tyr		
	170	175 180
Ala Phe Ala Phe Gly Tyr Met Asn Leu Phe Gly Val Ile Met Gln		
	185	190 195
Asn Val Phe Pro Ala Ala Pro Pro Trp Tyr Lys Ile Leu Tyr Gly		
	200	205 210
Leu Gln Ser Ala Asn Tyr Asp Met His Gly Ser Pro Gly Gly Leu		
	215	220 225
Ala Arg Ile Asp Lys Leu Leu Gly Ile Asn Met Tyr Thr Thr Cys		
	230	235 240
Phe Ser Asn Ser Ser Val Ile Phe Gly Ala Phe Pro Ser Leu His		
	245	250 255

Ser Gly Cys Ala Thr Met Glu Ala Leu Phe Phe Cys Tyr Cys Phe		
	260	270
Pro Lys Leu Lys Pro Leu Phe Ile Ala Tyr Val Cys Trp Leu Trp		
	275	285
Trp Ser Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Met		
	290	300
Ala Gly Ser Val Leu Ser Tyr Val Ile Phe Gln Tyr Thr Lys Tyr		
	305	315
Thr His Leu Pro Ile Val Asp Thr Ser Leu Phe Cys Arg Trp Ser		
	320	330
Tyr Thr Ser Ile Glu Lys Tyr Asp Ile Ser Lys Ser Asp Pro Leu		
	335	345
Ala Ala Asp Ser Asn Asp Ile Glu Ser Val Pro Leu Ser Asn Leu		
	350	360
Glu Leu Asp Phe Asp Leu Asn Met Thr Asp Glu Pro Ser Val Ser		
	365	375
Pro Ser Leu Phe Asp Gly Ser Thr Ser Val Ser Arg Ser Ser Ala		
	380	390
Thr Ser Ile Thr Ser Leu Gly Val Lys Arg Ala		
	395	400

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1206 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```

atggcaaacc ctttttcgag atggtttcta tcagagagac ctccaaactg ccatgtagcc    60
gatttagaaa caagtttaga tcccatcaa acgttggtga aggtgcaaaa atacaaaccc    120
68

```

```

gctttaagcg actgggtgca ttacatcttc ttgggatcca tcatgctggt tgtgttcatt 180
actaatcccc caccttggat cttcaagatc cttttttatt gtttcttggg cactttattc 240
atcattccag ctacgtcaca gtttttcttc aatgccttgc ccatcctaac atgggtggcg 300
ctgtatttca cttcatcgta ctttccagat gaccgcaggc ctcctattac tgtcaaagtg 360
ttaccagcgg tggaaacaat tttatacggc gacaatttaa gtgatattct tgcaacatcg 420
acgaattcct ttttggacat tttagcatgg ttaccgtacg gactatttca ttttggggcc 480
ccatttgtcg ttgctgccat cttattcgta tttgggtccac caactgtttt gcaaggttat 540
gcttttgcatt ttgggtatat gaacctgttt ggtgttatca tgcaaaatgt ctttccagcc 600
gctcccccat ggtataaaaat tctctatgga ttgcaatcag ccaactatga tatgcatggc 660
tcgcctgggtg gattagctag aattgataag ctactcggtg ttaatatgta tactacatgt 720
ttttcaaatt cctccgtcat tttcgggtgt tttccttcac tgcatcctgg gtgtgctact 780
atggaagccc tgtttttctg ttattgtttt ccaaaattga agcccttggt tattgcttat 840
gtttgctggt tatgggtggt aactatgtat ctgacacacc attattttgt agaccttatg 900
gcaggttctg tgctgtcata cgttattttc cagtacacaa agtacacaca tttaccaatt 960
gtagatacat ctcttttttg cagatgggtc tacacttcaa ttgagaaata cgatatatca 1020
aagagtgatc cattgggtgc agattcaaac gatatcgaaa gtgtcccttt gtccaacttg 1080
gaacttgact ttgatcttaa tatgactgat gaaccagtg taagcccttc gttatttgat 1140
ggatctactt ctgtttctcg ttcgtccgcc acgtctataa cgtcactagg tgtaaagagg 1200
gcttaa 1206

```

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1206 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```

atggcaaacc ctttttcgag atgggttcta tcagagagac ctccaaactg ccatgtagcc 60
gatttagaaa caagtttaga tccccatcaa acgttggtga aggtgcaaaa atacaaaccc 120
gctttaagcg actgggtgca ttacatcttc ttgggatcca tcatgctggt tgtgttcatt 180

```

actaatcccg caccttggat cttcaagatc cttttttatt gtttcttggg cactttattc 240  
 atcattccag ctacgtcaca gtttttcttc aatgccttgc ccacctaac atgggtggcg 300  
 ctgtatttca cttcatcgta ctttccagat gaccgcaggc ctcctattac tgtcaaagtg 360  
 ttaccagcgg tggaacaat tttatacggc gacaatttaa gtgatattct tgcaacatcg 420  
 acgaattcct ttttggacat tttagcatgg ttaccgtacg gactatttca ttatggggcc 480  
 ccatttgtcg ttgctgccat cttattcgta tttgggtccac caactgtttt gcaagggttat 540  
 gcttttgcag ttgggttatat gaacctgttt ggtgttatca tgcaaaatgt ctttccagcc 600  
 gctcccccag ggtataaaat tctctatgga ttgcaatcag ccaactatga tatgcatggc 660  
 tcgcctgggtg gattagctag aattgataag ctactcggta ttaatatgta tactacatgt 720  
 ttttcaaatt cctccgtcat tttcgggtgct tttccttcac tgcattccgg gtgtgctact 780  
 atggaagccc tgtttttctg ttattgtttt ccaaaattga agcccttggt tattgcttat 840  
 gtttgctggg tatgggtggc aactatgtat ctgacacacc attattttgt agaccttatg 900  
 gcaggttctg tgctgtcata cgttattttc cagtacacaa agtacacaca tttaccaatt 960  
 gtagatacat ctcttttttg cagatgggtca tacacttcaa ttgagaaata cgatatatca 1020  
 aagagtgatc cattggctgc agattcaaac gatatcgaaa gtgtcccttt gtccaacttg 1080  
 gaacttgact ttgatcttaa tatgactgat gaaccagtg taagcccttc gttatttgat 1140  
 ggatctactt ctgtttctcg ttcgtccgcc acgtctataa cgtcactagg tgtaaagagg 1200  
 gcttaa 1206

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1206 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

atggcaaacc ctttttcgag atggtttcta tcagagagac ctccaaactg ccatgtagcc 60  
 gatttagaaa caagtttaga tccccatcaa acgttggtga aggtgcaaaa atacaaaccc 120  
 gctttaagcg actgggtgca ttacatcttc ttgggatcca tcatgctgtt tgtgttcatt 180  
 actaatcccg caccttggat cttcaagatc cttttttatt gtttcttggg cactttattc 240

```

atcattccag ctacgtcaca gtttttcttc aatgccttgc ccatcctaac atgggtggcg 300
ctgtatttca cttcatcgta ctttccagat gaccgcaggc ctcctattac tgtcaaagtg 360
ttaccagcgg tggaaacaat tttatacggc gacaatttaa gtgatattct tgcaacatcg 420
acgaattcct ttttggacat tttagcatgg ttaccgtacg gactatttca ttatggggcc 480
ccatttgtcg ttgctgccat cttattcgta tttggtccac caactgtttt gcaaggttat 540
gcttttgcag ttggttatat gaacctgttt ggtgttatca tgcaaaatgt ctttccagcc 600
gtcccccat ggtataaaat tctctatgga ttgcaatcag ccaactatga tatgcatggc 660
tcgcctggtg gattagctag aattgataag ctactcggta ttaatatgta tactacagct 720
ttttcaaatt cctccgtcat tttcgggtgct tttccttcac tgcattccgg gtgtgctact 780
atggaagccc tgtttttctg ttattgtttt ccaaaattga agcccttggt tattgcttat 840
gtttgctggt tatggtgggc aactatgtat ctgacacacc attattttgt agaccttatg 900
gcaggttctg tgctgtcata cgttattttc cagtacacaa agtacacaca tttaccaatt 960
gtagatacat ctcttttttg cagatgggtca tacacttcaa ttgagaaata cgatataatca 1020
aagagtgatc cattggctgc agattcaaac gatatcgaaa gtgtcccttt gtccaacttg 1080
gaacttgact ttgatcttaa tatgactgat gaaccagtg taagcccttc gttatttgat 1140
ggatctactt ctgtttctcg ttcgtccgcc acgtctataa cgtcactagg tgtaaagagg 1200
gcttaa 1206

```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1206 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

```

atggcaaacc ctttttcgag atggtttcta tcagagagac ctccaaactg ccatgtagcc 60
gatttagaaa caagtttaga tccccatcaa acgttggtga aggtgcaaaa atacaaaccc 120
gctttaagcg actgggtgca ttacatcttc ttgggatcca tcatgctgtt tgtgttcatt 180
actaatcccg caccttggat cttcaagatc cttttttatt gtttcttggg cactttattc 240
atcattccag ctacgtcaca gtttttcttc aatgccttgc ccatcctaac atgggtggcg 300

```

```

ctgtatttca cttcatcgta ctttccagat gaccgcaggc ctcctattac tgtcaaagtg 360
ttaccagcgg tggaaacaat tttatacggc gacaatttaa gtgatattct tgcaacatcg 420
acgaattcct ttttggacat tttagcatgg ttaccgtacg gactatttca ttttggggcc 480
ccatttgtcg ttgctgccat cttattcgta tttgggtccac caactgtttt gcaagggttat 540
gcttttgcag ttggttatat gaacctgttt ggtgttatca tgcaaaatgt ctttccagcc 600
gctcccccag ggtataaaaat tctctatgga ttgcaatcag ccaactatga tatgcatggc 660
tcgcctgggt gattagctag aattgataag ctactcggta ttaatatgta tactacagct 720
ttttcaaatt cctccgtcat tttcgggtgct tttccttcac tgcatcctcg gtgtgctact 780
atggaagccc tgtttttctg ttattgtttt ccaaaattga agcccttggt tattgcttat 840
gtttgctggt tatggtgggt aactatgtat ctgacacacc attattttgt agaccttatg 900
gcaggttctg tgctgtcata cgttattttc cagtacacaa agtacacaca tttaccaatt 960
gtagatacat ctcttttttg cagatgggtca tacacttcaa ttgagaaata cgatatatca 1020
aagagtgatc cattgggtgc agattcaaac gatatcgaaa gtgtcccttt gtccaacttg 1080
gaacttgact ttgatcttaa tatgactgat gaaccagtg taagcccttc gttatttgat 1140
ggatctactt ctgtttctcg ttcgtccgcc acgtctataa cgtcactagg tgtaaagagg 1200
gcttaa 1206

```

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid (synthetic DNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

```

aatatgtata ctacatgttt ttcaaattcc 30

```

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleic acid



- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

gttaactcga gaaagtgccc atcagtgttc 30

(2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 bases
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

gttaacggta ccagaggaaa gaataacgc 29